
On the Obscuration of the Growing Supermassive Black Hole Population

Johannes Buchner



München 2015

On the Obscuration of the Growing Supermassive Black Hole Population

Johannes Buchner

Dissertation
an der Fakultät für Physik
der Ludwig-Maximilians-Universität
München

vorgelegt von
Johannes Buchner
aus Oberndorf bei Salzburg

München, den 27.02.2015

Erstgutachter: Prof. Dr. Kirpal Nandra
Zweitgutachter: Prof. Dr. Simon White
Tag der mündlichen Prüfung: 14.04.2015

1 Zusammenfassung

Aktive Galaxienkerne (AGN) werden durch das Wachstum super-schwere schwarze Löcher, die im Zentrum jeder massiven Galaxie sitzen, betrieben. Da enge Korrelationen ihrer Massen zu Eigenschaften der elliptischen Galaxienkomponente beobachtet werden, und durch ihre extreme Leuchtkraft ist es naheliegend, dass AGN einen wichtigen Baustein von Galaxien bilden. Der erste Schritt, AGN zu verstehen ist es, ihre Häufigkeit zu ermitteln, sowie die Leuchtkraft der Population. Dieses Unterfangen wird dadurch erschwert, dass die meisten AGN von Gas und Staub umgeben sind. Selbst im energiereichen Röntgenbereich, der in dieser Arbeit verwendet wird, wird die intrinsische Strahlung durch Absorption um mehrere Größenordnung verringert.

Die vorliegenden Doktorarbeit untersucht zuerst die Eigenschaften dieser Wolken, im speziellen ihre Geometrie, Säulendichteverteilung und ihr Verhältnis zur Leuchtkraft des AGN. Dazu werden ~ 300 AGN von der längst-beobachteten Röntgenregion, der Chandra Deep Field South Kampagne verwendet. Eine neue Bayesische Methode zur Spektralanalyse wurde entwickelt, um verschiedene physikalisch motivierte Modelle für den Aufbau der Wolken zu vergleichen. Das Röntgenspektrum reagiert, hauptsächlich dank Compton-Streuung, auf die Gesamtbedeckung der Quelle durch das Gas. Eine detaillierte Analyse zeigt, dass die Wolken mit einer Torus (“Donut”) Form konsistent sind, und sowohl vollständige Bedeckung als auch eine Scheiben-artige Konfiguration ausgeschlossen werden können. Außerdem ist eine weiteren Komponente höherer Dichte notwendig um zusätzlich beobachtete Compton-Reflektion zu erklären. Dies deutet auf eine strukturierte Formation hin, wie etwa ein Torus mit einem Dichtegradienten.

Die Untersuchung der gesamten AGN-Population inklusive der AGN mit hohen Säulendichten, verlangt eine große Stichprobe mit einem genauen Verständnis für die Stichprobenverzerrung, sowie fortgeschrittene statistische Inferenzmethoden. Diese Arbeit baut auf eine ~ 2000 AGN große Stichprobe die durch Röntgenemission detektiert wurde, bestehend aus mehrschichtigen Kampagnen aus den CDFS, AEGIS-XD, COSMOS and XMM-XXL Regionen. Die Röntgenspektren wurden im Detail mit einem physikalischen Spektralmodell analysiert, um die intrinsische Leuchtkraft, Rotverschiebung, sowie Säulendichte (N_H) für jedes Objekt zu erhalten, inklusive der Messunsicherheit. Außerdem wurden in dieser Arbeit neue statistische Methoden entwickelt um die richtige Assoziation zu optischen/infraroten Objekten zu finden, und um die Unsicherheiten durch Objekte ohne Pendant, der Rotverschiebungsmessung, sowie der Poissonfehler des Röntgenspektrums in alle Ergebnisse einzubinden.

Einen weiteren wichtigen Beitrag bildet eine Bayesische, nicht-parametrische Methode um die unverzerrte Dichte von AGN in kosmologischen Volumen als Funktion von intrinsischer Leuchtkraft, Rotverschiebung und Säulendichte (N_H) der verbergenden Wolken

zu rekonstruieren. Obwohl in dieser Methode lediglich Glattheit verwendet wird, kann dieser Ansatz dieselben Formen der Leuchtkraftverteilung sowie ihre Entwicklung rekonstruieren, die sonst oft in empirischen Modellen verwendet werden, jedoch ohne diese a-priori anzunehmen. Im Großen und Ganzen kann die Leuchtkraftverteilung, in allen Rotverschiebungsschalen, als Potenzgesetz mit einem Umbruchspunkt beschrieben werden. Sowohl die Normalisation als auch der Leuchtkraftumbruchspunkt entwickeln sich über den Lauf des Universums, allerdings zeigen die Daten keine Belege für eine Veränderung der Form der Verteilung. Dies deutet darauf hin, im Gegensatz zu Aussagen vorherigen Studien, dass der Rückkopplungsmechanismus zwischen AGN und beherbergender Galaxie immer gleich funktioniert, und sich nur die Anzahl und Größe der wachsenden Systeme verändert.

Die nicht-parametrische Rekonstruktionsmethode verwendet keine Annahmen darüber wie sich z.B. die Häufigkeiten von Säulendichte des verdeckenden Gases mit Leuchtkraft oder Rotverschiebung verändert. Dies erlaubt sehr robuste Schlüsse über den Anteil der verdeckten AGN ($N_H > 10^{22} \text{cm}^{-2}$), die $77_{-5}^{+4}\%$ der Population ausmachen sowie den Anteil der Compton-dicken AGN ($38_{-7}^{+8}\%$), die sich hinter enormen Säulendichten ($N_H > 10^{24} \text{cm}^{-2}$) verbergen. Insbesondere dass der letztere Anteil bestimmt werden konnte, lässt endlich Schlüsse darauf zu, wieviel AGN “verdeckt” wachsen. Außerdem suggeriert es, dass der Torus einen großen Teil des AGN verdeckt. Basierend auf der Leuchtkraft der gesamten AGN Population wurde die Masse, die über den Lauf der Zeit in schwarzen Löchern gesperrt wurde, geschätzt, und die Massendichte der supermassereichen schwarzen Löcher im heutigen Universum vorhergesagt.

Die Rekonstruktion bringt außerdem zu Tage, dass der Anteil der verdeckten AGN (insbesondere der Compton-dünnen AGN) eine negative Leuchtkraftabhängigkeit aufweist, und dass sich diese Abhängigkeit über die Geschichte des Universums entwickelt hat. Dieses Resultat wird in dieser Arbeit im Zusammenhang mit bestehenden Modellen interpretiert und ist möglicherweise ein Nebeneffekt eines nicht-hierarchischen Wachstums von AGN.

2 Abstract

Active Galactic Nuclei (AGN) are powered by the growth of super-massive black holes (SMBHs), which can be found at the centre of every massive galaxy. Due to tight scaling relationships of their masses with properties of their host spheroidal components, as well as the massive energy output AGN release, they are thought to play an important role in the formation and evolution of galaxies. The first step to understanding AGN is to determine their prevalence in the Universe, as well as the luminosity output of their entire population. This enterprise is hampered by the fact that most AGN are obscured by thick layers of gas and dust, making them difficult to detect. Even in the energetic X-ray wavelengths employed in this work, the intrinsic radiation of obscured AGN is suppressed by multiple orders of magnitude.

In this work I first study the properties of this obscurer, specifically its geometry, column density distribution and its relation to the AGN luminosity. For this, ~ 300 AGN from the deepest X-ray field to date, the Chandra Deep Field South survey, are used. I apply a novel Bayesian spectral analysis methodology to distinguish between several physically motivated models for the obscurer. The X-ray spectrum is, mainly due to Compton scattering, sensitive to the covering fraction of the obscurer. A detailed spectral analysis shows that the obscurer is consistent with a torus (“donut”) shape, but complete covering as well as disk-like configurations can be excluded. Furthermore, a high-density component is necessary to explain additional observed Compton-reflection beyond that expected from the line-of-sight obscuration, indicating a structured obscurer such as a torus with a density gradient.

The study of the population of AGN requires a large sample with detailed understanding of the selection effect and sophisticated inference techniques. A X-ray selected sample of ~ 2000 AGN from a multi-tiered survey including the CDFS, AEGIS-XD, COSMOS and XMM-XXL fields is analysed in detail. Through Bayesian spectral analysis with a physical model, the intrinsic luminosity, redshift and column density (N_H) is obtained for each source, including their uncertainties. This thesis also develops advanced statistical methodology for choosing the correct counterpart, and propagates the uncertainty from missing counterparts, redshift estimation as well as the Poisson noise from X-ray spectra into all final results.

Another important new contribution is a Bayesian non-parametric technique to reconstruct the unbiased number density of AGN in cosmological volumes as a function of intrinsic luminosity, redshift and column density (N_H). Despite only assuming smoothness, this approach is capable of reproducing the shapes commonly assumed for the luminosity function and its evolution, without assuming them a priori. Overall, the luminosity function appears to be consistent with a double powerlaw at all redshifts studied. Both the

normalisation and break luminosity evolve over time, while there is no evidence that the shape changes. This indicates that contrary to previous claims, the feedback mechanism works the same across the history of the Universe, but only the number and luminosity scale of the accreting systems changes.

The non-parametric reconstruction allows the study of the fraction of obscured AGN up to the Compton-thick regime in a very robust way, i.e. without assuming a luminosity or redshift-dependent behaviour a priori. About $77^{+4}_{-5}\%$ of AGN are obscured ($N_H > 10^{22}\text{cm}^{-2}$), while $38^{+8}_{-7}\%$ belong to the heavily obscured, elusive Compton-thick class ($N_H > 10^{24}\text{cm}^{-2}$). The latter fraction in particular finally constrains the importance of obscured growth phases in the life of accreting SMBHs. Based on the total luminosity output of the AGN population, the mass locked into black holes over cosmic time is estimated, and the mass density of relic SMBHs in the local Universe is predicted, and matches local estimates.

The large fraction of obscured AGN suggests that the obscuring torus must have a large angular extent. The non-parametric reconstruction also finds and characterises a negative luminosity dependence for the fraction of obscured AGN, in particular those that are Compton-thin, which are less prevalent at high luminosities. Additionally, this luminosity dependence appears to evolve with redshift. These findings are discussed in the context of existing models and it is concluded that the observed evolution may be to first order a side-effect of a anti-hierarchical growth of super-massive black holes.

Contents

1	Zusammenfassung	1
2	Abstract	3
3	Introduction	12
3.1	A brief historical overview of AGN research	12
3.2	The structure of AGN	13
3.2.1	Accretion disk	13
3.2.2	Corona	14
3.2.3	Obscuring material in the vicinity of the SMBH	14
3.2.4	Soft scattering component	18
3.2.5	Soft excess	18
3.3	The life of AGN and their host galaxies	18
3.3.1	Selection methods	21
3.3.2	Census of the AGN population	23
3.3.3	Completing the demographics including Compton-thick AGN	27
3.3.4	Relation with host galaxies	29
3.4	Outline of this thesis	31
4	Data	32
4.1	X-ray telescopes	32
4.1.1	Chandra	32
4.1.2	XMM-Newton	32
4.1.3	Survey fields	33
4.2	Detection of X-ray point sources	36
4.2.1	A detailed description for the CDFS	36
4.2.2	Reduction of the other survey fields	39
4.3	Extraction of spectra	42
4.4	Background model	42
4.4.1	Background model definition	43
4.4.2	Goodness of Fit and Model discovery	43
4.4.3	Using the background model	46
4.5	Association with optical/infrared counterparts	48
4.6	Bayesian association method (NWAY)	49
4.6.1	Distance-based matching	49

4.6.2	Magnitudes	51
4.6.3	Auto-calibration	52
4.6.4	Implementation	52
4.7	Redshift estimation	53
4.7.1	Robust photometric redshift probability distributions	55
4.7.2	X-ray spectroscopic redshifts (XZ)	59
4.8	Sample selection	64
4.9	Sub-sample for study of the obscurer geometry	66
5	The nature of the obscurer around SMBH	67
5.1	Model definitions	69
5.2	Statistical analysis methods	73
5.2.1	Parameter estimation	73
5.2.2	Model comparison	75
5.2.3	Model verification	76
5.2.4	Priors	77
5.2.5	Implementation	78
5.3	Results	78
5.3.1	Propagation of redshift uncertainty	79
5.3.2	Source 179: An example of Bayesian model comparison	81
5.3.3	Model selection on the full sample	85
5.4	Discussion	89
5.4.1	X-ray spectral analysis methodology	89
5.4.2	The nature of the obscurer	92
5.5	Conclusions	95
6	The relationship between obscurer and accretion luminosity	100
6.1	Methodology	103
6.1.1	Luminosity function analysis	104
6.1.2	Non-parametric approach	110
6.2	Results	111
6.2.1	Obscured and Compton-thick fractions	111
6.2.2	Obscuration-dependent evolution	116
6.2.3	Luminosity-dependence and evolution of the obscured fraction . . .	119
6.2.4	Evolution of Compton-thick AGN	120
6.2.5	N_H distribution	121
6.3	Discussion	123
6.3.1	The role of Compton-thick AGN	123
6.3.2	Obscuration-dependent evolution	125
6.3.3	Luminosity-dependence and evolution of the obscured fraction . . .	125
6.4	Conclusions	132

7	Accretion history of the Universe	134
7.1	Shape of the X-ray Luminosity function	134
7.1.1	Introduction	134
7.2	Observed luminosity function	136
7.2.1	Evolution of the luminosity function	138
7.2.2	Comparison with previous works	138
7.3	The accretion history of the Universe	142
7.3.1	Inferring infalling matter from X-ray radiation	142
7.3.2	The importance of accretion in obscured black holes	144
7.4	Discussion	148
7.4.1	The shape and evolution of the luminosity function	148
7.4.2	Accretion history of the Universe	149
7.5	Conclusions	151
8	Summary	152
9	Acknowledgements	157

List of Figures

3.2.1 Illustration of the typical shapes of the spectral features	15
3.2.2 Illustration of the Unification paradigm	15
3.2.3 Illustration of the main components in the central engine of AGN	19
3.2.4 Illustration of the main components in the central engine of AGN under a clumpy torus model	19
4.1.1 Observing fields in the sky	34
4.2.1 Typical response function	36
4.2.2 Area curve	38
4.2.3 Observation field sizes	40
4.4.1 Comparison of the background data from Source 318 with background models.	45
4.4.2 Comparison of the background data from Source 179 with background models.	47
4.7.1 Redshift distribution of the full sample	56
4.7.2 Redshift distribution per field	56
4.7.3 Photo-z error analysis.	57
4.7.4 Demonstration of the kernel convolution of the PDZ.	58
4.7.5 Redshift constraints obtained using the XZ method.	61
4.7.6 Continuation of Figure 4.7.5	62
4.7.7 Dependence of information gain (IG) on column density and number of X- ray counts.	63
4.7.8 Evaluation of XZ redshift estimates.	64
4.8.1 Luminosity-redshift plot of the full sample	65
5.1.1 Geometry of models	69
5.1.2 Illustration of model spectra	71
5.1.3 Model hierarchy	74
5.3.1 Degenerate solutions in the parameter space	79
5.3.2 Parameter estimation with and without propagating redshift uncertainties	80
5.3.3 Observed spectrum and convolved model of object 179	82
5.3.4 Marginalised parameters	83
5.3.5 Model comparison preferences of the sample	86
5.5.1 Inferred possible geometries	95
5.3.6 Luminosity-redshift plot of the sample	96
5.3.7 Evidence contribution from each source with secure spectroscopic redshift .	97
5.3.8 Parameter distributions	98
5.3.9 Comparison of the derived column density and luminosity	99

List of Figures

6.1.1 Broken powerlaw luminosity function model	105
6.1.2 Illustration of the two smoothness priors used	110
6.2.1 Redshift evolution of the space density of AGN	115
6.2.2 Redshift evolution of the space density of AGN split by obscuration	116
6.2.3 Luminosity-dependence of the obscurer	117
6.2.4 Evolution of the obscured fraction	118
6.2.5 The evolution of the density of Compton-thick AGN	121
6.2.6 Column density distributions at various redshifts	122
6.3.1 Illustration of the population-averaged obscuring toroid	126
6.3.2 Four models that address the luminosity-dependence of the obscured fraction	127
7.1.1 Illustration of a bending powerlaw distribution	135
7.2.1 The non-parametric luminosity function at various redshift bins.	137
7.2.2 The evolution of luminosity function parameters	139
7.2.3 Total luminosity function	140
7.3.1 Conversion between X-ray luminosity, bolometric luminosity and accretion rate	143
7.3.2 Evolution of the X-ray luminosity density	145
7.3.3 Evolution of the accretion density of AGN	146
7.3.4 Black hole mass density	147
8.0.1 The survey properties of the eROSITA mission	154
8.0.2 Flux-Number prediction for the eROSITA all-sky survey	155

List of Tables

3.1	Comparison of methodologies in luminosity function works	25
4.1	Survey fields	33
4.2	Sample statistics for the individual data extraction steps.	35
4.3	Background model parameters for two sources.	44
5.1	Parameters of the models used.	70
5.2	Model selection results for object 179	84
5.3	Sample model comparison	87
5.4	Sub-sample model comparison	88
5.5	Catalogue (excerpt)	90
5.6	Compton-thick AGN catalogue	91
6.1	Key statistics on the fraction of obscured AGN.	114
6.2	Predictions of the discussed models.	131

List of Algorithms

6.1	Stan code for estimating the field	112
6.2	Stan code for estimating the field (cont.)	113

3 Introduction

3.1 A brief historical overview of AGN research

Active Galactic Nuclei (AGN) were first noticed in 1908 by Edward A. Fath as peculiar objects that appeared on photographic plates like stars but showed spectra similar to galaxies (“nebulae” at that time). Several instances were found where the centres of galaxies showed strong emission lines, which were first systematically collected by Carl K. Seyfert in 1943. In radio, bright point-like sources were found, which were first thought of as galactic stars. Martin Schmidt recognised in 1960 that one of these quasi-stellar radio sources (quasars) was at redshift 0.158 based on identification of a redshifted Balmer emission line series. This placed it at extreme distances, which was very controversial at the time. A number of papers then associated them with distant galaxies, finally recognising them as a distinct class of objects.

Since this point a vast number of these objects have been found, and it was found that they are always placed at the centre of galaxies (the active nucleus of galaxies). Some of them, like the quasar 3C273 or Cyg A, outshine their host galaxy in luminosity (e.g. Baade & Minkowski 1954). Furthermore, the emission from these sources was particularly variable – within days the luminosity could vary by several orders of magnitude. If this emission was from stellar light, it should originate from several million stars like our suns (Woltjer 1959), and their emission would have to be co-ordinated on a time-scale of days. Light speed dictates that such a system must not exceed the size of the solar system. A hundred million stars packed in the solar system would immediately collapse into a black hole¹, according to the predictions of the relatively new theory of general relativity. However, this conundrum essentially gave the solution at the same time: If every year, a star from the galaxy would be added to that black hole, the gravitational energy released is comparable to the luminosity needed. This idea formed the basics of our current understanding of AGN as radiation originating via accretion onto black holes (Salpeter 1964; Zel’dovich & Novikov 1964; Lynden-Bell 1969). Since the discovery of AGN, our understanding of their nature has been substantially improved by observational contributions from different wavelengths as well as theoretical work. Below, I lay out our current of the AGN emission processes, in particular those relevant for the X-ray wavelengths.

¹The term “black hole” had still to wait until 1967 for its invention by John Wheeler.

3.2 The structure of AGN

3.2.1 Accretion disk

The theoretical foundation for understanding the conversion of accreted matter into radiation and black hole growth was developed in Pringle & Rees (1972) and Shakura & Sunyaev (1973). An accretion disk forms around a supermassive black hole (SMBH) as gas arrives from galactic, parsec-scale environments to distances such as Sun-Earth or Sun-Neptune scales with non-zero angular momentum. Heat is produced by turbulence and magnetic field line re-connection. The black-body radiation of the innermost, hottest part of the accretion disk peaks in the ultra-violet (UV) for AGN, while outer rings peak at subsequently lower energies, following a powerlaw distribution in temperature. Such a powerlaw composite of black-body emitters yields a powerlaw spectrum at wavelengths shorter than UV.

In the formalism of Shakura & Sunyaev (1973), created to explain spectra of accreting stellar black holes, the luminosity output of accretion disk is a fraction μ (radiative efficiency) of the accreted rest mass energy, i.e.:

$$L_{bol} = \mu c^2 \cdot \dot{M}. \quad (3.2.1)$$

Here, \dot{M} denotes the accretion rate, c is the speed of light while L_{bol} refers to the luminosity output over the entire electromagnetic spectrum (bolometric luminosity). The remainder of the accreted rest mass goes into growth of the black hole. Typical values for the radiative efficiency include $\mu = 6\%$ for the accretion disks around non-rotating black holes. Naively, one would expect that infalling matter would spin up black holes, at which point the radiative efficiency can reach $\mu = 42\%$. However, due to random orientation of the infalling matter from large scales, and frame-dragging effects in the accretion disk, the black hole spin may not necessarily become maximal over long periods of time (King & Pringle 2006; King et al. 2008; Wang et al. 2009).

The high radiative efficiencies make AGN the most efficient source of radiation known in the Universe (short of antimatter-matter recombination). However, if the system is starved of infalling material (low accretion rates), advection is the dominant process (Narayan & Yi 1994, 1995a,b; Blandford & Begelman 1999), and the efficiency can be substantially lower. On the other hand if the accretion rate is very high, radiation pressure becomes important and can prohibit all infall. The maximum (average) luminosity a system can thus attain, the Eddington luminosity, is derived by equating the gravitational pull to the radiation pressure due to Thomson scattering off electrons, and assuming spherical symmetry. This yields

$$L_{Edd} \cong 1.26 \times 10^{44} \left(\frac{M_{BH}}{10^6 M_{\odot}} \right) \text{erg/s} \quad (3.2.2)$$

where M_{BH} refers to the mass providing the gravitational field (the supermassive black hole) and M_{\odot} is the solar mass. The AGN luminosity relative to the Eddington limit is referred to as the Eddington rate.

3.2.2 Corona

Substantial AGN samples with good X-ray spectral resolution were obtained using the *HEAO-1* (Mushotzky 1984), *Einstein* (Reichert et al. 1985), *EXOSAT* (Turner & Pounds 1989) and *Ginga* (Nandra & Pounds 1994) satellites. These works find that the shape of the intrinsic X-ray spectrum of AGN can be described to first order as a power law $F(E) = C \cdot E^{-\Gamma}$ (see Figure 3.2.1). Nandra & Pounds (1994) observed values for the photon index Γ range between 1.4 and 2.8, and found that its distribution in local AGN can be approximated by a Gaussian of mean 1.95 and standard deviation 0.15.

To explain the X-ray spectrum, Katz (1976) proposed that a hot electron gas, the corona, exists near the accretion disk. Simple analytic models by Sunyaev & Titarchuk (1980, 1985) have successfully reproduced powerlaw-like shapes. The fast electrons increase the energy of the incident spectrum (a powerlaw, see above) by inverse Compton scattering. This model also predicts that at some energy, determined by the highest electron velocities, the X-ray powerlaw truncates. The observed spectrum up to the hardest X-rays can be well-described by the empirical formula $F(E) = C \cdot E^{-\Gamma} \exp \{-E/E_c\}$. Zdziarski et al. (2000) found that the corona model matches observations obtained with the *Compton Gamma Ray Observatory*, *Ginga*, and *EXOSAT* satellites. The steep cut-off at energies around $E_c \sim 100$ keV is found to vary from object to object. This Comptonization temperature of the corona depends on the cooling power of the seed photons (the UV photons from the accretion disk). However, it also depends on the geometry and optical depth of the corona, which are as of yet unknown.

Long-term observations with *EXOSAT* show that the X-ray spectrum undergoes variability on timescales of days and shorter (McHardy 1989). Ghisellini et al. (1993) argue that this supports the accretion disk / corona picture: In that model the seed photon population propagates its fast variability into the X-ray spectrum. If instead a plasma was emitting a thermal spectrum on its own, temperature changes of the corona would necessarily be slow. Nevertheless, the process responsible for maintaining the corona remains unclear.

3.2.3 Obscuring material in the vicinity of the SMBH

The intrinsic power law is often observed to be modified by cold material in the line of sight (Mushotzky 1982; Turner et al. 1997a; Risaliti et al. 1999). The most important physical effects of such an obscuring screen is photo-electric absorption, which reduces the number of low-energy photons (illustrated in Figure 3.2.1).

Why are some AGN obscured, while in others, the intrinsic radiation can be observed directly? The simplest explanation is that they are intrinsically similar, but different viewing angles give rise to the observed differences (Unification paradigm, Antonucci 1982, 1993). Under this hypothesis, the simplest common obscuring geometry is a toroidal structure usually referred to as “the torus”. Figure 3.2.2 illustrates this geometry. Originally, this idea was developed to explain the radio morphology of AGN (Antonucci 1982; Antonucci & Miller 1985; Antonucci 1993), as well as why some AGN show broad emission lines (type 1,

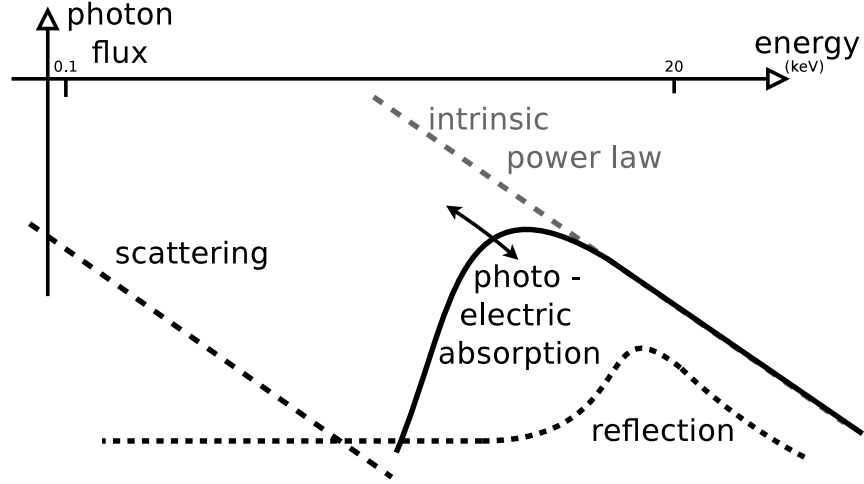


Figure 3.2.1: Illustration of the typical shapes of the discussed spectral features. Emission lines and absorption edges have been omitted for simplicity.

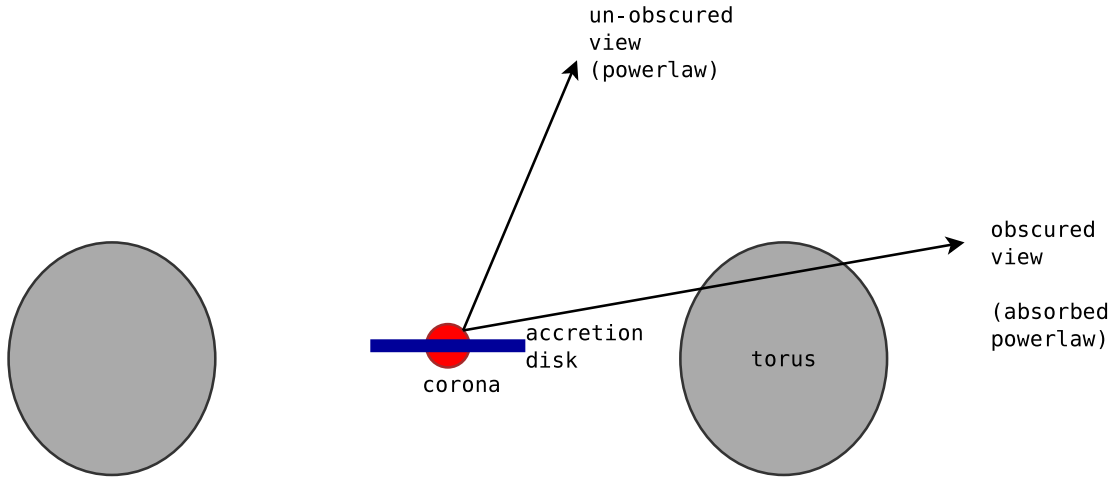


Figure 3.2.2: Illustration of the Unification paradigm. Depending on the viewing angle, the intrinsic radiation (powerlaw) is seen directly or obscured by a toroidal obscurer.

see Seyfert 1943; Schmidt & Green 1983; Khachikian & Weedman 1974) while others show only the narrow core of emission lines (type 2, obscured line-of-sight to a central broad line region). Antonucci & Miller (1985) targeted a ionised gas cloud in the vicinity of a type 2 AGN and observed its spectrum in polarised light. The observations showed broad lines compatible with the spectrum of a type 1 AGN, which is interpreted to be mirrored by the ionised gas cloud. This led to increased confidence that indeed, type 2 AGN are intrinsically type 1 AGN, but an obscurer in the line-of-sight hides the broad emission lines. This has made to the Unification paradigm popular in explaining the majority of the diversity of AGN via viewing angles and an obscuring “torus”.

X-ray observations do not resolve AGN and spectral models can not directly distinguish the scales at which e.g. photo-electric absorption occurs. However, there are additional effects. The obscuring material in the vicinity of the SMBH can also Compton-scatter photons into or out of the line-of-sight (LOS). Compton scattering, unlike photo-electric absorption, is anisotropic. Seen from a single, unresolved viewpoint the geometry of the obscurer influences the observed spectrum. For example, compared to a sphere enclosing the X-ray source, a torus geometry with a certain opening angle produces a smaller Compton reflection hump between $\sim 10 - 30$ keV as well as different line emission strengths (see Figure 3.2.1), related to the solid angle illuminated by the source and the surface area exposed to the observer (Murphy & Yaqoob 2009; Brightman & Nandra 2011a). X-ray fluorescent line emission, in particular from Iron (Fe) which has a relatively high abundance and a high yield for Fe $K\alpha$ (6.4keV), is also important. The relative strength of the Fe K feature is influenced by Compton scattering and has thus been investigated as an indicator on the geometry (e.g. Matt et al. 1996, 1997; Brightman & Nandra 2011a). Awaki et al. (1991) find Fe $K\alpha$ equivalent widths, at least in some obscured AGN observed with *Ginga*, to be inconsistent with a uniform obscurer around the source, but compatible with a torus geometry.

In the strongest interpretation of the Unification paradigm all AGN share the same obscuring geometry. Then the fraction of the sky covered directly corresponds, via random sampling of the viewing angles, to the fraction of obscured AGN (of the order of 50%). It was understood early on that such extended, cold, molecular material would tend to collapse and form stars (Krolik & Begelman 1988). The torus was thus suggested to be made up from small, cold clumps, as illustrated in Figure 3.2.4. Pier & Krolik (1992) have argued that a extended torus can be continuously maintained by radiation pressure if the torus is clumpy. Another strong argument for clumpiness comes from the not uncommon phenomenon that AGN vary in their classification between “unobscured” and “obscured” and back over the years (e.g. Turner & Pounds 1989; Risaliti et al. 2002; Markowitz et al. 2014). Monitoring NGC 1365, Risaliti et al. (2007) found variability in the level of obscuration within hours. They then model a cloud on a Keplerian orbit occulting the X-ray source, which subsequently needs to be closely associated with the X-ray source (closer than a few hundred Schwarzschild radii). As mentioned above, the obscuring material is observed to be cold and molecular. Gaskell et al. (2008) thus argue that the obscurer must also either be distant (parsec-scale) or self-shielding (e.g. through a transition between accretion disk and torus).

Recently, high-resolution mid-infrared interferometry by Bartscher et al. (2013) identified parsec-scale dust components in 16 nearby AGN. Also, the central 100 parsec region of host galaxies shows substantial gas structures in Hubble Space Telescope (HST) images (e.g. Ferrarese et al. 1996; van der Marel & van den Bosch 1998). Guainazzi et al. (2005) shows that the presence of such structures correlates with Compton-thin X-ray obscuration. A recent HST survey of nearby AGN by Prieto et al. (2014) indicates that such structures are common and indeed comparable to the structure in the Milky Way Galactic centre as mapped by Herschel (Molinari et al. 2011). Maiolino & Rieke (1995) show that many nearby AGN which are intermediate between the type-1 and type-2 classification are observed through dust lanes and edge-on host galaxy viewing angles, raising suspicion that obscuration contributed by the host galaxy affects the classification. All of these works indicate that a galaxy contribution may not be negligible. A two-component obscurer model was proposed by Matt (2000), also to explain the correlation between type-2 AGN and host galaxy morphology, star formation and dust content. In summary, the X-ray obscurer may be complex and a combination of layers on multiple scales which leave their cumulatively imprint on the intrinsic emission. Other components not mentioned above have been detected in some AGN, which include Compton reflection off the accretion disk (and relativistically blurred $K\alpha$ line emission), several components which can cause warm absorption (e.g. broad/narrow line regions, relativistic bulk outflows, disk winds) and X-ray emission from radio jets. The typical spectral quality at the redshifts of this work however do not justify the inclusion of such effects. Since only cold absorption is considered in this work, the derived obscuring column densities should be considered as the “effective” absorption, combining the effects of warm and cold absorption from all components on various scales.

This work defines as obscured AGN those with LOS neutral hydrogen equivalent column densities of $N_H > 10^{22}\text{cm}^{-2}$; those with lower values are called unobscured AGN. The most heavily obscured sources known fall into the Compton-thick regime, defined via the inverse of the Thomson cross section, i.e. $N_H \geq \sigma_T^{-1} = 1.5 \times 10^{24}\text{cm}^{-2}$. For simplicity many studies adopt the border at $N_H = 10^{24}\text{cm}^{-2}$, as is done in this work. At such high column densities, the spectrum is dominated by Compton-scattering through the obscuring screen and typically shows a very flat spectrum. The amount of Compton-scattering depends on the geometry of the obscurer.

In the local Universe, about half of all type-2 AGN are obscured by Compton-thick matter (Maiolino & Rieke 1995; Bassani et al. 1999). The spectra of five nearby Compton-thick AGN, such as NGC 1068, have been investigated by Matt et al. (2000). They also find strong infrared emission from the obscurer. However, they also find that the infrared emission is dominated by star burst emission in several of their sources. The study of such nearby galaxies with high count statistics have been used to verify the validity of X-ray spectral models for Compton-thick sources, such as those made available by Brightman & Nandra (2011a) and Murphy & Yaqoob (2009), which include the physical effects named above in toroidal geometries irradiated by a central source with a powerlaw spectrum. Using a variety of missions, recent results e.g. by Arévalo et al. (2014) on the Circinus galaxy and Bauer et al. (2014) on NGC 1068 indicate that these models are a good description,

but may require the inclusion of additional Compton reflection (before the obscurer is encountered).

3.2.4 Soft scattering component

Obscured AGN often show an excess of photons at soft energies (see Figure 3.2.1). This component can be parameterised by an additional powerlaw with a normalisation that is a few percent of the intrinsic (unobscured) powerlaw (e.g. Reichert et al. 1985; Turner et al. 1997b; Guainazzi et al. 2005). One simplistic interpretation is that a fraction of the intrinsic radiation leaks through the obscurer without being Compton scattered or photo-electrically absorbed (“partial covering”). More physical explanations for this component include Thomson scattering off ionised material within the torus opening angle, which reflects the intrinsic spectrum past the obscurer into the line of sight (Krolik & Kriss 1995; Reynolds & Fabian 1995). This picture is illustrated in Figure 3.2.3. In some AGN like NGC 1068, a large ionisation cone is observed in optical high-resolution images (Antonucci & Miller 1985). In this obscured AGN it has been shown that the polarised optical spectrum looks like that of an unobscured AGN, indicating that the reflection works as a mirroring screen that allows an indirect view of the inner region. Ogle et al. (2003) employ detailed observations of this region with *Chandra* high-resolution X-ray spectroscopy. They find emission lines by ionised species in an optically thin gas, confirming the scenario of a warm mirror. Such ionised regions have been speculated by Reynolds & Fabian (1995) to be outflows of the AGN (illustrated in Figure 3.2.4) or heated circumnuclear gas. This component is referred to as “scattering” or “soft scattering” in this work, as opposed to “Compton scattering” or “reflection” via the cold obscurer.

3.2.5 Soft excess

In unobscured AGN, extrapolating the 2 – 10 keV power law to softer energies shows an excess of soft X-rays in some sources (Turner & Pounds 1989; Gondhalekar et al. 1997). Among the processes that have been proposed to explain this soft component are a second corona with lower temperature (Czerny & Elvis 1987; Bechtold et al. 1987), a photo-ionized reflection spectrum by the accretion disk (Ross & Fabian 1993), or ionised absorption by a relativistically blurred disk wind (Gierliński & Done 2004). In medium-to-high redshifts, where the bulk of the objects of this work’s interest lie, the soft excess component is outside the observed energy range. For this reason, the soft excess is not considered in this work.

3.3 The life of AGN and their host galaxies

AGN are interesting objects in their own right since they allow study of physical processes under extreme conditions. However, there are other research fields for which AGN are important. Here three such areas are named briefly, which are discussed in more detail in the following sections.

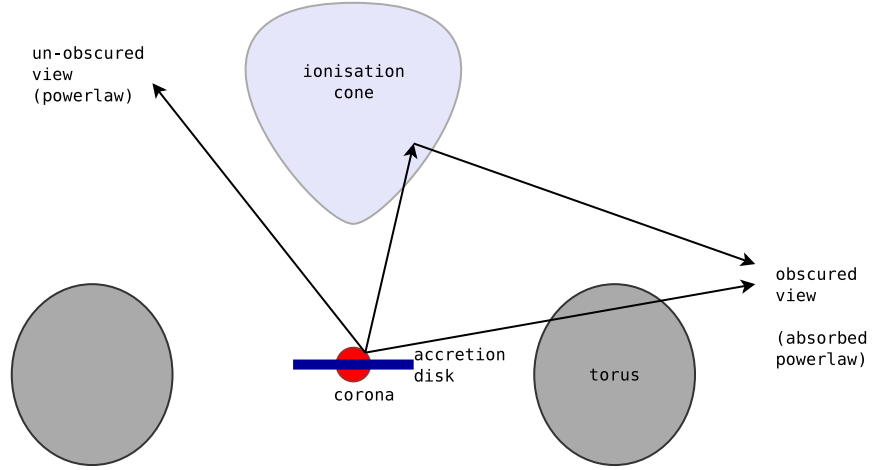


Figure 3.2.3: Illustration of the main components in the central engine of AGN. The accretion disk (dark blue), the corona (red), the X-ray obscurer (torus, grey) and ionised clouds in the opening angle of the torus providing Thomson scattering (light grey).

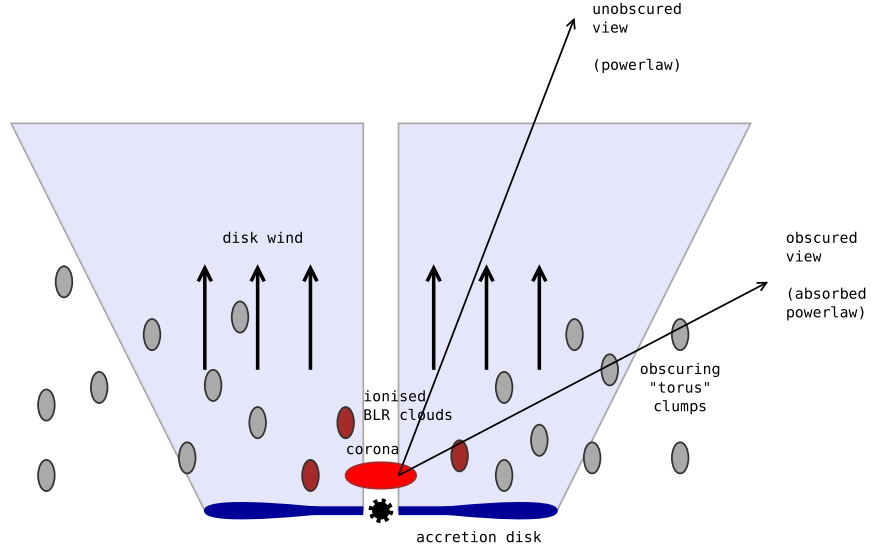


Figure 3.2.4: Another possible scenario for the central engine of AGN. Here, the accretion disk (dark blue) is shown truncated by the last stable orbit around the black hole. The corona (red) is located above. Here, the “torus” is an arrangement of clumps of cold gas, which are vertically supported by the disk wind. The innermost clouds are partially ionised (“broad-line region”).

Compact massive objects with masses up to $10^{10}M_{\odot}$ have been found in nearby galaxies by their gravitational influence on the stellar population and gas in the innermost galactic regions (see Kormendy & Ho 2013, for a recent review). Since these enormous masses are concentrated in small regions (smaller than solar system size), these have been interpreted to be supermassive black holes (see Genzel et al. 2010, for detailed exclusion of alternative hypotheses). Based on the fact that AGN are powered by accretion of matter onto black holes, Soltan (1982) was the first to predict the mass density of such relics by the cumulative luminosity output of the AGN population over cosmic time. The luminosity output at any specific time can be characterised using the luminosity function, which is the number density of AGN per luminosity interval per comoving volume element ($\phi(L)/dL/dV_c$, measured in number/(erg/s)/Mpc³). Based on scaling relationships between properties of the galaxy elliptical components (see Section 3.3.4) and a compilation of local galaxy luminosity functions, Marconi et al. (2004) estimated the local black hole mass density to be $\rho_{BH} = 4.6^{+1.9}_{-1.4} \times 10^{14}M_{\odot}\text{Gpc}^{-3}$. The comparison of such estimates for the relic black holes masses with the luminosity output of AGN, however requires a good understanding of the luminosity function of the entire AGN population, including those with obscured lines of sights.

AGN are detected only in a small fraction of galaxies. Under the hypothesis that the triggering of AGN is dependent on the environment their host galaxy provides, numerous studies have investigated AGN as a phase in the life of galaxies. Sanders et al. (1988) suggested that galaxy mergers make gas available to star formation and also triggers feeding of the SMBH, leading to the appearance of the host galaxy as luminous infrared galaxies first and later as bright quasars. This idea became popular due to further development by Hopkins et al. (2005a), who predict that mergers lead to a heavily obscured SMBH accretion phase followed by a short, luminous AGN episode. AGN have also been suggested to have considerable influence on their host galaxy e.g. in quenching star formation (such ideas are discussed more below in Section 3.3.4). To constrain such feedback processes, it is necessary to develop a better understanding of the correlation of AGN properties, such as their black hole masses, Eddington accretion rates and obscuration, and galaxy properties, such as their star formation rate and whether they have experienced a merger (morphology). One approach to understanding the impact of possible AGN feedback processes is to measure the frequency of the AGN phenomenon as a function of various galaxy properties (e.g. morphological type, specific star formation rate). To understand the strength of feedback mechanisms, if thought to be related to the accretion process, for the formation of galaxies it is again important to understand the luminosity function of AGN.

Cosmic X-ray background (XRB) radiation has first been detected by Giacconi et al. (1962) with Geiger counters on an Aerobee rocket. The *HEAO-1* and *ROSAT* missions show the XRB to be largely isotropic over the sky, with a broad peak in the spectrum around 12keV, but spanning the 0.1 – 1000keV range (Fabian & Barcons 1992). The availability of missions with higher angular resolution has resolved a substantial fraction of the soft (Brandt & Hasinger 2005), and a smaller fraction of the hard X-ray emission (Worsley et al. 2005) into individual AGN, indicating that the dominant contributor to the diffuse XRB is the AGN population. These works suggest that the remainder of the

XRB emission, especially the 20keV peak where only a tiny fraction of individual AGN are detected, is due to a heavily obscured AGN sub-population which elude detection. Population synthesis models (e.g. Gilli et al. 2007, more are discussed below) have attempted to reconcile the XRB spectrum with the known AGN population from luminosity function studies. Such studies also require an accurate measurement of the evolution of the AGN luminosity function over cosmic time and a distribution of the obscuring column in AGN.

3.3.1 Selection methods

Early works, such as Schmidt (1968), used UV/optical selection to determine the luminosity function. This selection uses the tail of the UV black body radiation of the accretion disk and is capable of detecting AGN out to high redshifts. However, criticism has been raised against this selection as obscuration hides the accretion disk, leaving out about 80% of the population (Maiolino & Rieke 1995). Additionally, ongoing star formation can result in UV emission and contaminate the sample.

Baldwin et al. (1981) suggested an alternative selection based on the two emission line ratios of Balmer lines and ionised Oxygen/Nitrogen (BPT diagram). This method uses the response of gas to the extreme ionisation power of high-energy photons from AGN to separate out lower ionisation processes (e.g. star formation) in passive galaxies (i.e. those without AGN activity) from AGN. This selection circumvents the issue of accretion disk obscuration, because the emission region of narrow lines is believed to lie at larger distances than the obscuring torus. The first spectra of a large magnitude-limited sample of 486 galaxies was analysed by Ho et al. (1997). Due to their proximity, spectra could be taken in the nuclear region. They subtracted the stellar light and determined the strengths and widths of emission lines to identify AGN. Such studies laid an anchor point to constrain the local AGN number density and luminosity function (Schmidt & Green 1983). Technical advances of these studies led to the Sloan survey which detected tens of thousands of AGN (Kauffmann et al. 2003). However, as spectra of larger distances are obtained, severe contamination by the host galaxy is unavoidable, especially for fainter AGN which then may remain undetected (“host dilution”). Also, such samples often include galaxies with lower luminosity emission lines typically of lower degrees of ionisation (low-ionisation nuclear emission-line regions, LINERs). This class of galaxies were first detected and defined by Heckman (1980), who argued that they constitute the low-luminosity end of the AGN population. Today, evidence exists that LINERs are not AGN. Yan & Blanton (2012) determine the contribution to the line emission from various radii centered at the center of galaxies. Based on assuming a gas distribution in galaxies, they find that the line emission does not decrease outwards as expected from a central, AGN ionisation source, but that the ionisation source is distributed throughout the galaxy. This result was later confirmed by Sarzi et al. (2010), who use integral field unit observations (spatially resolved spectra) of nearby galaxies. They argue that stars are the more likely source for excitation. Also, they note that detailed analysis shows that several of their galaxies which are classified by BPT diagrams to be LINERs (and not AGN), are in fact AGN. The separation between AGN and passive galaxies is thus not straightforward due to a mixture of processes.

In X-ray, the luminosity of passive galaxies is observed to be 1-2 orders of magnitude below those of typical AGN. Massive compact binary systems and hot interstellar gas in particular in star-forming galaxies can contribute some soft emission below 1keV. This radiation does however not exceed X-ray luminosities of $L_X = 10^{41}$ erg/s (Ptak et al. 1999; Norman et al. 2004). For this reason, the threshold $L_X > 10^{42}$ erg/s is commonly used for a clean separation between AGN and star-forming, passive galaxies. This luminosity cut also forms the definition of AGN as used in this thesis. Lower luminosity AGN do not contribute strongly to the total luminosity of the population as relevant to e.g. the total growth of supermassive black holes or the XRB. X-ray selection can thus address the issue of host dilution effectively and derives clean AGN samples.

Since X-rays can penetrate the obscurer, their selection function is simpler and only depends on the properties of the AGN, namely instantaneous X-ray luminosity, obscuration and redshift. The bias specific to X-ray selection is, to first order, determined by photo-electric absorption. This effect is dependent only on the column density of neutral hydrogen in the line-of-sight. This selection bias can be well modelled because it is largely independent of the potentially complex gas phase and geometry. The selection bias against Compton-thick AGN is severe, but can also be well quantified if the spectra of AGN, i.e. effects beyond photo-electric absorption, are understood. The spectra of a large sample of $z = 0.5 - 4$ AGN are discussed in Chapter 5 and tested against various models.

At very high obscuring columns, X-rays do not penetrate the obscurer efficiently. Several authors have suggested selection criteria based on infrared colour-magnitude diagrams (Lacy et al. 2004; Donley et al. 2012; Stern et al. 2012; Assef et al. 2013). These selection criteria exploit that the re-radiated, thermal emission from the torus shows a powerlaw-like spectral shape due to the combination of black-body radiators over a range of temperatures. These selection criteria have successfully detected a large number of obscured AGN. However, the study of local Compton-thick AGN by Matt et al. (2000) demonstrated that often, the infrared excess is dominated by star formation in the host galaxies. In these selections, the separation space between star-forming, passive galaxies is narrow, comparable to BPT diagrams. Another issue is that the efficiency of this selection varies between unobscured, mildly obscured and heavily obscured AGN in a non-uniform, luminosity-dependent way that is difficult to quantify. While these methods have been used to identify Compton-thick candidates (see Section 3.3.3), this selection is thus limited in its use to determine a complete census of AGN.

A demographic study of the entire distribution of obscuring line-of-sight column densities also allows another insight into the processes creating and maintaining the obscurer. A significant fraction of obscured AGN implies a large geometric extent of a distant obscurer. For instance, in a strictly geometric picture of a torus following the Unification paradigm, the fraction of obscured AGN can be translated into a viewing angle. Models for various obscuring processes can then be restricted to reproduce such column density distribution. Probing the obscurer in such a way is only possible through X-ray studies, which yield the column density of the obscuration in the line of sight and where the selection bias is known.

Finally, X-rays are produced very close to the accretion process, which suggests that

the X-ray luminosity may be a useful proxy (or part of a proxy) for the instantaneous accretion rate. The benefits of X-ray selection have thus motivated a large body of studies to investigate the luminosity function and obscuration of AGN using X-ray telescopes.

3.3.2 Census of the AGN population

Substantial technological developments had to be achieved before the first determinations of X-ray luminosity functions. Some of the first substantial samples were obtained around 1990 using e.g. the *Einstein* (Maccacaro et al. 1991) and *ROSAT* (Boyle et al. 1993) satellites. These works characterised the hard (2 – 10keV) and soft (0.5 – 2keV) luminosity respectively as observed (i.e. not corrected for absorption). However, already at that time it was suspected (e.g. Lawrence & Elvis 1982) that the obscured fraction is luminosity-dependent. The introduction of Chapter 6 discusses the luminosity and redshift dependence of the obscured fraction in more detail.

A major advance came from Ueda et al. (2003), who collected large hard-band detected samples from the *HEAO-1*, *ASCA* and *Chandra* missions. To enlighten the discussion of luminosity functions, I digress for a moment into the methodologies of X-ray luminosity functions and describe their methodology in some detail. The X-ray detections have been associated to optical counterparts using images from follow-up observations at the positions of X-ray detections. Spectra, or photometric redshift methods, have been used to identify the objects and estimate their redshift. This results in a redshift completeness ranging from 85% – 100% depending on the field and X-ray instrument; X-ray detections without redshift are not used further (except for up-scaling the final number density). For each source, they computed a ratio of the detected counts in the soft and hard band, (hardness ratio, $HR = (\text{hard} - \text{soft})/(\text{hard} + \text{soft})$). Using the total count rate (CR) and the hardness ratio, an estimate for the obscuring column density (N_H) and intrinsic luminosity was derived using a simple template model. In a subset of their sample, the full X-ray spectra also were available to determine these quantities. Such fits only converge for sources with a high number of counts. The successful spectral fits were used to determine the distribution of the column density (N_H function), and empirically find a simple functional form. Then, the selection bias both in luminosity and obscuration is taken into account when fitting the full data set with an empirical model for the hard, intrinsic luminosity function (evolutionary models are discussed in more detail in the introduction to Chapter 7).

Subsequent studies have attempted to improve on this methodology in several ways: Brusa et al. (2007) has developed advanced techniques to handle ambiguous associations with counterparts, which is necessary due the improved angular resolution of current (*Chandra* and *XMM-Newton*) X-ray telescopes. Salvato et al. (2009) have substantially improved the catastrophic outlier rate and accuracy by advancing photometric redshift methods via SED fitting of AGN/galaxy hybrid template spectra. Such advances have improved some later luminosity function determinations, but not all works adopted these techniques. Furthermore, La Franca et al. (2005) and Ebrero et al. (2009) improved the determination of the N_H function by fitting a larger number of X-ray spectra. This is important because the count rate and hardness ratio are not unique indicators of N_H and luminosity,

because the spectrum of an AGN has several components (see above). Instead of using a best fit point estimate for e.g. the luminosity, Aird et al. (2010) provided substantial improvements in the handling of uncertainties. They used a formalism that propagates the uncertainties from the detected counts, subsequently creating a probability distribution for the luminosity of each object. In that process, they also incorporate the uncertainty arising from photometric redshift methods, and also handle those X-ray detections without optical counterpart. However, the Aird et al. (2010) study does not consider source obscuration, and determines an absorbed, hard-band luminosity function. The studies of La Franca et al. (2005), Ebrero et al. (2009) and Ueda et al. (2014) also included Compton-thick AGN, but their exact fraction of the total population remain poorly constrained. Finally, Brightman & Nandra (2011a) developed advanced spectral models for AGN obscuration geometries which extend into the Compton-thick regime, and Ueda et al. (2014) has taken advantage of these models to improve both the description of the selection function and the spectral analysis.

Table 3.1 shows a comparison of the methodologies of X-ray luminosity function studies. So far, none of the previous studies has derived the intrinsic luminosity function of AGN while incorporating the uncertainties from count detectors and redshift determinations. Also, the use of hardness ratios has remained wide-spread. However, to constrain the fraction of Compton-thick AGN, detailed analysis of low-count spectra, with realistic spectral models, may be essential. This thesis contributes major advances in methodology. Chapter 5 develops the methodology for analysing spectra which produces probability distribution on the luminosity and obscuring column while propagating uncertainties in redshift. The new spectral analysis methodology can also robustly handle low count sources, freeing us from the use of hardness ratios. Based on the luminosity, redshift and obscuring column probability distributions from spectral analysis of a sample of ~ 2000 AGN (sample presented in Chapter 4), Chapter 6 develops a non-parametric technique to determine the joint N_H distribution and luminosity function simultaneously. The following results are particularly robust as the demographics of the AGN population is determined without relying on empirical models. In Chapter 6, the dependence of the determined obscured AGN fraction and Compton-thick AGN fraction on luminosity and redshift is investigated in detail. Chapter 7 presents the total luminosity function of AGN (including Compton-thick AGN). Section 7.3 presents predictions for the relic black hole mass density for the newly derived, total luminosity functions. Unlike previous predictions, in this work the fraction of Compton-thick AGN does not have to be assumed. Further minor advances include a new method for probabilistic association between X-ray detections and multiple optical/IR catalogues (Section 4.6) and incorporating the systematic uncertainty of photometric redshift estimators (Section 4.7.1) which is also propagated all the way through the subsequent analysis.

Optical and X-ray luminosity function works as well as other studies have yielded a large body of important results. Determinations of the X-ray luminosity function using the *Einstein* (Maccacaro et al. 1991) and *ROSAT* (Boyle et al. 1993) satellites find the local X-ray luminosity function to be, to first order, a broken powerlaw. The distribution is relatively flat up to a certain break luminosity L_* , above which it declines strongly. This

Reference	detection band	L from	N_H from	Uncertainty from counts	Poisson likelihood	Uncertainty from redshifts	intrinsic LF dist.	CT handled	model-independent verification	
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Boyle et al. (1993)	soft	counts	-	✓	✓	-				✓
Ueda et al. (2003)	hard	CR/HR	HR	✓	✓		✓	✓		
Hasinger et al. (2005)	soft	counts	-	✓	✓					✓
La Franca et al. (2005)	hard	mixed: CR/HR, spectra	HR or spectra				✓	✓	✓	
Silverman et al. (2008)	soft+hard	counts	-	✓	✓					✓
Ebrero et al. (2009)	various	mixed: CR/HR, spectra	HR or spectra				✓	✓	✓	✓
Yencho et al. (2009)	soft+hard	counts	-		✓					✓
Aird et al. (2010)	hard	counts	-	✓	✓	✓				
Ueda et al. (2014)	hard	HR	HR, some spectra		✓		✓	✓	✓	
This thesis	hard	spectra	spectra	✓	✓	✓	✓	✓	✓	✓

Table 3.1: Comparison of methodologies in luminosity function works. Column 1 shows the publication; Column 2 lists the band used for detection (soft: $0 - 2\text{keV}$, hard: $2 - 8\text{keV}$ or $2 - 10\text{keV}$); Column 3 and 4 show how the luminosity and obscuring column N_H were estimated; either from counts (absorbed LF), from counts and hardness ratios, or from a full spectral analysis. Column 5 shows whether the Poisson uncertainty arising in detected counts was propagated. Column 6 shows whether the work has followed a maximum likelihood approach based on Poisson draws (Marshall et al. 1983) in their parametric LF estimation; some works instead use a simple χ^2 comparison with the expected number of detections. Column 7 lists whether uncertainties from photometric redshift methods were incorporated in the analysis. Column 8 shows whether the derived luminosity function is in absorbed luminosity (hard or soft band) or in intrinsic luminosity. Column 9 indicates whether the work has used a distribution in N_H ; typically describing the luminosity-dependence of the fraction of obscured AGN, and its evolution with redshift. Column 10 lists whether Compton-thick AGN have been handled explicitly by the work. Column 11 indicates whether the work has incorporated an approach to independently verify the assumed empirical models by a independent non-parametric method (e.g. $1/V_{max}$ by Schmidt 1968).

is in agreement with optical studies (e.g. Schmidt 1968; Marshall et al. 1983). Section 7.1 discusses the shape of the luminosity function in more detail.

Luminosity function works both in optical/radio (e.g. Schmidt 1968; Schmidt & Green 1983) and X-ray (Maccacaro et al. 1991; Boyle et al. 1993) found that the AGN space density shows a strong increase from $z = 0$ to $z \approx 1$. Schmidt (1968) proposed that the evolution is due to an increase in the total number of AGN at all luminosities (Pure Density Evolution, PDE), where the shape of the luminosity function remains constant. Mathez (1976) proposed, alternatively, that the luminosity of the entire population may increase (Pure Luminosity Evolution, PLE), thereby increasing the number of detected AGN. Schmidt & Green (1983) construct the luminosity function from a sample of 90 AGN detected by their $\text{Ly}\alpha$ emission and find that PLE is the more appropriate model. The same conclusion is reached by Maccacaro et al. (1991) on the basis of 420 *Einstein* X-ray detected AGN, and Boyle et al. (1993) using 42 bright *ROSAT*-detected AGN. Subsequent works (e.g. Miyaji et al. 2000; Ueda et al. 2003) confirm this trend and show that around $z \approx 1$, the evolution stops and transitions into a broad plateau from $z = 1 - 3$, with a slight decline. The optical luminosity function of AGN (e.g. Richards et al. 2006b; Croom et al. 2009) shows the same behaviour. One consequence of this evolutionary behaviour is that more luminous systems show their peak in number density earlier than less luminous systems.

This suggest that the typical luminosity of AGN decreases over cosmic time, an effect that has been termed “downsizing” (Barger & Cowie 2005). One simplistic interpretation is that more massive systems (having higher Eddington limits) have provided the circumstances for accretion earlier in cosmic time, while small-scale systems accrete later. Such an interpretation is counter-intuitive in a Universe that grows by hierarchical collapse of structures and because any individual super-massive black hole can not decrease its mass. However, AGN have to be viewed in the context of their host galaxies’ growth over cosmic time. A similar downsizing is seen in galaxies with regard to their star formation histories (which peak later for more massive galaxies) (Cowie et al. 1996) and potentially several other properties (see discussion in Fontanot et al. 2009). Downsizing has been reproduced in some semi-analytic models (e.g. Fanidakis et al. 2012; Enoki et al. 2014; Hirschmann et al. 2014). In the model of Hirschmann et al. (2014) for instance, downsizing is explained by luminous AGN emptying the gas reservoir of their host galaxy by feedback processes, although star formation processes also play a role. In the simulation of Fanidakis et al. (2012), the co-evolution of AGN and host galaxies creates down-sizing via varying the relative importance of two different accretion channels. One channel is the accretion of hot gas from the halo, which is important throughout cosmic time to produce low-luminosity AGN (“radio mode”). Another important channel, activated by mergers, causes star bursts and accretion of cold gas onto black holes near the Eddington limit. This mode is particularly active at $z = 1 - 3$ and causes the brightest luminosities (“quasar mode”). When this channel loses importance towards low redshift, bright AGN disappear, and the peak of the luminosity distribution shifts to lower values.

Schmidt et al. (1995) investigated the optical luminosity function studies based on $\text{Ly}\alpha$ emission at $z > 3$ and finds the overall shape of the number density of AGN to show an

exponential decline with increasing redshift. This lack of sources had been noticed before: Osmer (1982) for instance searched for AGN at $z > 3$ via emission lines in an area of 5 deg^2 , expecting, given no comoving space density evolution, 10 detections with their sensitivity. Only weaker galaxy lines but no AGN (e.g. via $\text{Ly}\alpha$ emission lines) were detected. As a result, they argued that the space density must decrease towards high redshift. Such a decline was claimed by Silverman et al. (2008) but has remained controversial for some time in the X-ray. Initial results from soft (Hasinger et al. 2005) and hard (La Franca et al. 2005; Yenko et al. 2009) X-ray luminosity functions did not find such a trend. The reconstruction of the $z > 3$ luminosity function is difficult for X-ray observations due the decreased flux but also due to the uncertain association to optical counterparts in deep field observations of high resolution and systematic errors of photometric redshift methods. Through advances in counterpart association and redshift estimation methods as well as improved data quality by dedicated multi-wavelength follow-up campaigns, evidence for a decline in the number density at $z > 3$ has now also been convincingly claimed in the X-ray (Brusa et al. 2009; Civano et al. 2011; Vito et al. 2014; Ueda et al. 2014). It is worth emphasising that only the last work by Ueda et al. (2014) takes the obscuration of AGN into account and thus this trend still requires further confirmation.

It is noteworthy that the cosmic star formation history also shows a broad peak in the $z = 1 - 3$ range, as well as a decrease above $z = 4$ (see Madau & Dickinson 2014, for a recent review). Already Boyle & Terlevich (1998) have thus suggested that star formation of host galaxies and AGN are intimately linked. The dominant phase in the life of galaxies when gas is condensed into stars and accreted into black holes thus appears to be the period of $z = 0.5 - 4$ (a substantial amount of time passes between $z = 0$ and $z = 1$ which justifies inclusion down to 0.5). A complete census of the AGN population in this range is thus most important.

3.3.3 Completing the demographics including Compton-thick AGN

A striking result from X-ray surveys in the last 15 years was the high number of obscured AGN detected. This highlights the importance of considering obscuration for a full census of AGN. As mentioned above, this is required to understand the SMBH relics in the local Universe. To this end, the fraction of obscured AGN as a function of redshift and luminosity has been the subject of a number of studies (see Chapter 6 for a discussion). Selection effects make the two effects difficult to study in isolation. A further poorly understood subject is the fraction of the Compton-thick AGN, due to their elusive nature. Addressing these issues is vital for completing the demographic census of the AGN population.

Another indication that obscured AGN are an important part of the AGN population comes from the spectral shape of the XRB. Setti & Woltjer (1989) noted that the relatively flat ($\Gamma \approx 1.4$) spectral shape observed by the *HEAO-1* satellite can be accommodated with the intrinsic X-ray spectrum ($\Gamma \approx 1.9$) if a large fraction of AGN show substantial obscuration. Worsley et al. (2005) subtract the contribution of the known AGN population down to the deepest *Chandra* and *XMM-Newton* fields from the XRB. The spectral distribution of the missing component is consistent with heavily obscured AGN. In the XRB popula-

tion synthesis model of Akylas et al. (2012) up to 50% of the AGN population may be hidden behind Compton-thick line of sights. In these models, however, a level of Compton scattering due to the geometry has to be assumed to fit the spectrum, which is degenerate with the fraction of Compton-thick AGN. The popular model of Gilli et al. (2007) assumes a equal amount of Compton-thick and Compton-thin, obscured AGN (the latter already being more numerous than unobscured AGN). The XRB thus provides a good motivation to find Compton-thick AGN and to study the importance of Compton-thick reflection in the AGN population (due to geometry).

One obvious method to detect Compton-thick AGN would be to use higher energies which penetrate the obscurer. Burlon et al. (2011) compiled a three-year survey using 15 – 200keV detection on board the *Swift* satellite, and follow up these objects with X-ray observations in lower energies to determine their column densities. In their sample of 200 sources, only 9 are Compton-thick AGN. Based on the selection bias of their instrument, they estimate that the true number of Compton-thick AGN in their field may be five times higher. This illustrate how elusive Compton-thick AGN are even in high energy bands, in some parts due to the detector efficiency and powerlaw decrease towards high energies. In contrast, Brightman et al. (2014) have identified 100 Compton-thick AGN in three deep *Chandra* surveys. The large number is due to the higher sensitivity of X-ray detectors in lower energy regimes, and due to substantial high-energy intrinsic emission falling into the observing window, as many of these objects are at high redshift. They applied spectral fitting of data collected in the 0.5 – 10keV range (and hardness ratio techniques for low count sources). The same surveys are also part of this work.

To discover Compton-thick AGN for study, several techniques have been developed. One is to first detect type 2 AGN selected via optical narrow emission lines. This parent sample is then followed up by X-ray observations to constrain the obscuring column density distribution of X-ray obscuring columns therein. Risaliti et al. (1999) find that 50% of their compiled ~ 40 sources from the local Universe are Compton-thick AGN. Several authors have developed and applied flux ratio diagnostics (e.g. based on [OIII], [NeV]) for the detection of Compton-thick AGN (e.g. Gilli et al. 2010; Vignali et al. 2010; Jia et al. 2013; Mignoli et al. 2013). Overall, such techniques may allow the identification of Compton-thick AGN candidates for study, as well as the fraction of Compton-thick AGN in the obscured population, but, at least on its own, not the fraction of Compton-thick AGN to the total population or their space density. Furthermore, Brightman & Nandra (2011b) showed that the [OIII] to X-ray luminosity ratio does not constitute a reliable identification of Compton-thick AGN, due to substantial scatter in this ratio for Compton-thin and Compton-thick AGN. It is also worth pointing out that the two classifications methods (type 1/2 in optical, X-ray column densities) are not interchangeable, and may lead to contradictory classifications in up to 30% of objects (Merloni et al. 2014).

Another method to detect Compton-thick AGN is via the strong infrared emission of their dusty obscurer (infrared excess as compared to passive galaxies). Fiore et al. (2008) select a sample based on infrared excess and red optical colours in the *Chandra* Deep Field South (CDFS) field. They find that such sources show similar hardness ratios as *Chandra* X-ray detected, X-ray spectrally confirmed Compton-thick AGN. Fiore et al. (2008) and

Fiore et al. (2009) then estimate the space density of their infrared-selected “Compton-thick” AGN in two fields. Another method (BzK diagram) is used by Alexander et al. (2011) and verified by measuring X-ray column densities. However, while they find that many X-ray detected Compton-thick AGN show an infrared excess, they caution that star forming galaxies and unobscured AGN may severely contaminate infrared-excess selected sample. After employing X-ray information for verification of their sample, they derive a conservative space density. Chapter 6 compares the results reached in this work against those estimates.

This work attempts to construct a full census of the AGN population using hard-band X-ray selection. The ability to determine the fraction of Compton-thick AGN is improved in this thesis by significant methodological advances that take advantage of the full spectral information in the 0.5 – 10keV band and improved handling of the uncertainties.

3.3.4 Relation with host galaxies

The study of the AGN population (including obscured AGN) to understand the growth of SMBHs also has important implications for other research fields. Supermassive black holes are compact regions in the Universe where masses of $10^5\text{--}10^9 M_\odot$ (solar masses) appear within a few astronomical units (Sun-Earth distance). Such extreme physical environments can not be found in other objects or reproduced in laboratories, and thus provide a unique possibility to study the behaviour of matter and radiation under general relativistic effects. One might think that such massive, exotic objects are rare. The opposite is true: virtually every massive galaxy appears to have such an object in its centre (Richstone et al. 1998; Kormendy & Ho 2013). Furthermore, the mass of the black hole shows a tight correlation with the mass of the host galaxies bulge (Magorrian et al. 1998) as well as the luminosity-weighted line-of-sight velocity dispersion of the central region ($M_{BH} - \sigma$ relation, Gebhardt et al. 2000).

These relationships have been taken as a sign that SMBHs are an important component in the make-up of galaxies which co-evolves with their hosts (Ferrarese & Merritt 2000). A number of processes have been proposed that connect the relatively tiny sphere of influence of the black hole with its host galaxy. They range from loose interactions, such as just a shared gas reservoir connecting star formation and accretion rate, to more direct interactions. As mentioned above, a significant part of any accreted material is converted into radiation, which can yield luminosities exceeding the binding energy of the galaxy. A plausible interaction of this energy with the interstellar medium of the host galaxy could thus have a fundamental impact on the evolution of the galaxy (see Cattaneo et al. 2009). A more securely known mechanism is powerful, collimated jets extending beyond galaxy scales of some AGN, such as in the prominent case of M87, which have been unambiguously traced back to the supermassive black hole by high-resolution observations (Doeleman et al. 2012). Also, based on broad optical absorption lines, there is substantial evidence for fast ($v \approx 10^4 \text{ km/s}$) bulk outflows in bright AGN (see e.g. Turnshek 1988, and references therein). Motivated by the latter, Silk & Rees (1998) proposed that AGN outflows may sweep up gas, a model further expanded by King (2003), who compares the situation to a

momentum-driven stellar wind bubble. Gas is heated, up to a radius determined by the black hole mass, which effectively stops star formation inside and leaves behind a elliptical component. This process can reproduce the $M_{BH} - \sigma$ relation directly. Given that interactions are plausible, simulations in cosmological volumes have implemented various recipes for AGN triggering and AGN-galaxy interaction mechanisms. Croton et al. (2006) proposed a maintenance feedback mode, possibly in the form of jets, which can heat up or expel gas in massive galaxies, stopping and suppressing star formation. They invoke this process to reduce the otherwise overpredicted high end of the galaxy mass function as well as to reproduce the bimodality of the colour distribution of galaxies. Menci et al. (2008) explicitly models the blast-wave model of King (2003) in phases of high Eddington accretion rates to explain the observed downsizing of the AGN population as well as the luminosity-dependence of the obscured AGN fraction. Hopkins et al. (2005a) developed a semi-empirical model that assumes strong AGN feedback leading to the suppression of star formation. In their model, mergers of gas-rich galaxies lead to gas inflows towards the galactic centre, obscuring and feeding the SMBH. The available gas also triggers strong bursts in star formation. The SMBH grows and at some point becomes a bright AGN, sheds the gas and suppresses further star formation in the central region but also SMBH accretion. This leaves behind a normal, passive galaxy. Their model explains a wide range of AGN population properties, including the shape and evolution of the luminosity function, the obscuration distribution as well as its luminosity-dependence. In recent years, following Croton et al. (2006), it has also become common practice to include AGN feedback in semi-analytic models (Somerville et al. 2008; Fanidakis et al. 2012; Hirschmann et al. 2012) to jointly explain the population properties of galaxies and AGN.

These models have been very useful in making specific predictions, with which one can learn about the triggering mechanism, accretion mechanism and its relation to star formation rate. For instance, many of these models trigger AGN activity when galaxies are merging. One testable prediction is then that merging galaxies, which appear disturbed in optical images, should be associated with AGN presence, as traced by X-ray emission. A number of studies have employed asymmetry and concentration measures (e.g. Grogin et al. 2003, 2005; Pierce et al. 2007; Gabor et al. 2009; Kocevski et al. 2011) as well as visual classification (Kocevski et al. 2011) to compare the appearance of AGN hosts to passive galaxies. These studies do not find any enhanced sign of distortion due to previous major mergers episodes in their morphologies, and thus argue against major merger as the main mode of triggering AGN, as employed in the model of Hopkins et al. (2006a) (first proposed by Sanders et al. 1988). Instead, minor mergers (which may leave as-of-yet undetected traces) as well as secular processes such as disk fragmentation (e.g. models by Bournaud et al. 2007, 2011) are thought to be more important. Another proposal of these models is that AGN activity should be associated with enhanced star-formation activity. Rosario et al. (2011) investigated the optical colour distribution of AGN host galaxies identified by X-ray activity and found no appreciable differences to passive galaxies with the same stellar mass. However, inside star-forming galaxies, as detected by *Herschel*, AGN activity is more common (Rosario et al. 2013) than in non-starforming galaxies. Hickox et al. (2014) argued that the direct connection between level of star-formation and AGN

luminosity may still be present at long timescales, but due to the high variability of AGN is not detectable in individual galaxies, but only on the average.

3.4 Outline of this thesis

Understanding the components of the AGN machinery, and their interaction, is a long-standing puzzle, with important ramification for other fields in physics and astronomy. This thesis investigates the role of gas and dust in the near vicinity of accreting supermassive black holes and creates a census of the AGN demographics in luminosity and obscuration over cosmic time.

Chapter 4 describes the data used in this study and how they were treated.

In Chapter 5, the structure of the absorbing material in AGN is investigated using X-ray spectroscopy. Multiple physically models for the obscurer, as well as the presence of other AGN components, are tested in the $z \approx 0.5 - 4$ sample.

Chapter 6 undertakes a detailed population study of AGN, and derives their distribution in luminosity, redshift and obscuring column. In Section 6.2.1, the fraction of Compton-thick AGN averaged over cosmic time is constrained, as well as the fraction of all obscured AGN. That chapter then focusses on investigating the relation between the obscured AGN fraction and the intrinsic luminosity, as well as evolutionary trends. Several processes for the maintenance and evolution of the obscurer are discussed in Section 6.3.

Chapter 7 presents the shape of the luminosity function and its evolution. The evolution of the number density of the AGN population is derived. In Section 7.3, the total luminosity output of the AGN population is used to constrain the corresponding growth onto supermassive black holes and predict their relic mass density.

Finally, the findings are summarised in Chapter 8. Predictions for future missions and followup work are presented.

4 Data

4.1 X-ray telescopes

This research used data collected by X-ray telescopes on board two satellites, the “Chandra X-ray Observatory” (Chandra; Weisskopf et al. 2000) and the “X-ray Multi-Mirror Mission - Newton” (XMM-Newton; Jansen et al. 2001). Both telescopes were launched into orbit in 1999. Compared to previous missions (e.g. ROSAT, Ginga, ASCA), Chandra and XMM-Newton deliver substantial improvements in terms of angular resolution, spectral resolution and sensitivity. Both use grazing incident mirrors of Wolter type I design to focus X-ray photons. This type of optics can deliver small point spread functions (PSFs) on-axis. Off-axis however the sensitivity decreases and the PSF becomes substantially larger and non-uniform in shape. Both telescopes are designed for observations in the 0.5 – 10keV energy range.

4.1.1 Chandra

Chandra uses four nested mirrors with a focal length of 10 meters. The Chandra mirrors achieve a PSF of 0.5 arcsec Full Width Half Maximum (FWHM) on-axis.

Chandra is equipped with two detectors: The High-Resolution Camera (HRC) is optimised for high-resolution spectroscopy while the Advanced CCD imaging spectrometer (ACIS; Garmire et al. 2003) is designed for imaging spectroscopy. The latter has been extensively used to conduct deep X-ray survey programs. ACIS consists of two CCD arrays, ACIS-I (2x2) and ACIS-S (1x6). ACIS-S is meant for receiving spectra of the on-board High Energy Transmission Grating (HETG) and also features two back-illuminated CCDs, which provide a wider energy band and slightly higher spectral resolution than front-illuminated CCDs. ACIS-I consists solely of front-illuminated CCDs which have a lower background noise and thus are suitable for long, low count exposures. The pixel size is 0.492 arcsec.

The main benefit of Chandra is its high resolution and small PSF. This allows distant sources in very deep surveys to still be identified. For this reason, Chandra/ACIS has been used to conduct the deepest X-ray survey to date. The effective area is twice that of ROSAT, about 400cm² at 5 keV.

4.1.2 XMM-Newton

XMM-Newton is equipped with three telescopes, each made of 58 nested mirrors with a focal length of 7.5 meters. The PSF of XMM-Newton is five arcsec FWHM on-axis.

Table 4.1: Survey fields

Survey	CDFS	AEGIS-XD	C-COSMOS	XMM-XXL
Survey area	464 arcmin ²	1010 arcmin ²	0.9 deg ²	20 deg ²
Total/central exposure time	4 Ms	2.4 Ms/800 ks	1.8 Ms/160 ks	10 ks

Two of the telescopes split their beams with one half going to a grating spectrometer and the second half being received by two Metal Oxide Semiconductor CCD detectors (MOS 1/2; Turner et al. 2001). Both MOS 1 and 2 consist of seven front-illuminated chips, but are rotated with respect to each other. This ensures that the gaps between chips of one detector are covered by the other detector. The pixel size is one arcsec.

The remaining third mirror is dedicated to a single CCD array (PN; Strüder et al. 2001). It consists of 12 back-illuminated chips with a pixel size of four arcsec. The back-illuminated chips feature a high quantum efficiency compared to the front-illuminated chips.

XMM-Newton provides a substantially larger effective area as compared to Chandra (1000cm²). For survey work, both MOS and PN can be used simultaneously, and together form the European Photon Imaging Camera (EPIC).

4.1.3 Survey fields

Among the numerous surveys the Chandra and XMM-Newton missions have carried out, this work uses data from four regions of the sky (see Figure 4.1.1): the Chandra Deep Field South (CDFS, Xue et al. 2011), the All Wavelength Extended Groth strip International Survey (AEGIS, Davis et al. 2007), the Cosmological evolution Survey (COSMOS, Scoville et al. 2007) from Chandra, and the XMM-XXL survey (PI: Pierre) observed with XMM-Newton. The subsequent analysis focuses on the region of the CDFS which is covered by the 4Ms Chandra observations and the part of the AEGIS field which has been surveyed by Chandra for a total of 800 ks (AEGIS-XD, Nandra et al., submitted). In the COSMOS field the region covered by the Chandra observations performed between November 2006 and June 2007 (C-COSMOS, Elvis et al. 2009) was used. For XMM-XXL, the equatorial subregion is used. Table 4.1 presents information on the individual X-ray fields used in this paper.

I now describe the data handling of these surveys in more detail, starting with the Chandra Deep Field South. The data reduction and source detection followed Laird et al. (2009) and is published in Rangel et al. (2013b,a) (performed by Cyprian Rangel). The other surveys, where processing is very similar, are described afterwards in Section 4.2.2.

A cosmology of $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $\Omega_M = 0.3$ and $\Omega_\Lambda = 0.7$ is adopted. Solar abundances are assumed. The galactic photo-electric absorption along the line of sight direction is modelled with $N_H \approx 8.8 \times 10^{19}$, 1.3×10^{20} , 2.7×10^{20} and $2.2 \times 10^{21} \text{ cm}^{-2}$ for the CDFS, AEGIS-XD, COSMOS and XMM-XXL fields respectively (Dickey & Lockman 1990; Stark et al. 1992). In this work, luminosity (L) always refers to the intrinsic (absorption-corrected) luminosity in the 2 – 10 keV rest-frame energy range.

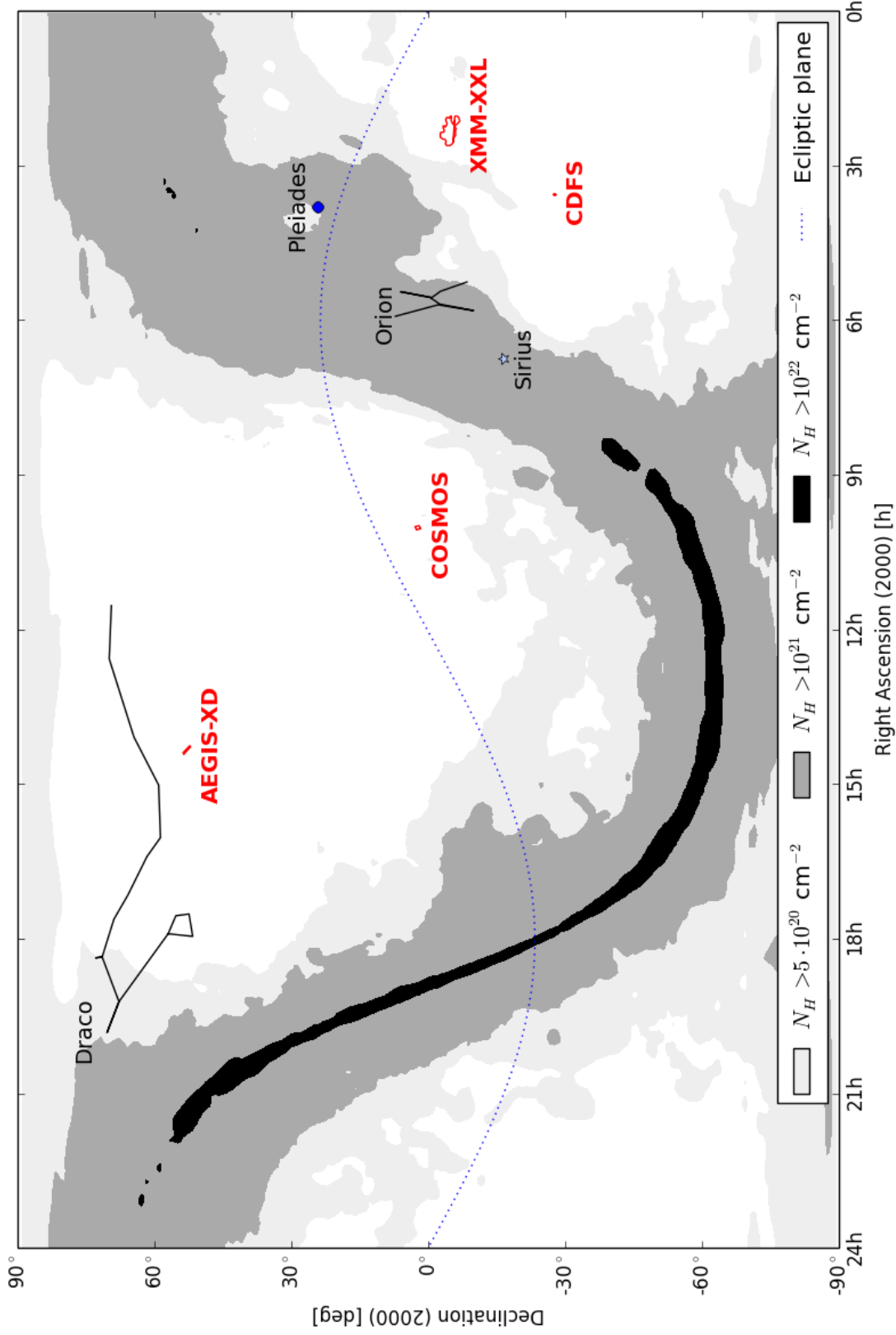


Figure 4.1.1: Observing fields used in this study on the sky. The CDFS, AEGIS-XD, COSMOS and XMM-XXL survey fields are outlined in red. Additionally, I a galactic absorption map (Dickey & Lockman 1990) is overlaid and the ecliptic plane and a few well-known objects are shown for orientation.

Table 4.2: Sample statistics for the individual data extraction steps.

Section	Extraction step	CDFS	AEGIS-XD	C-COSMOS	XMM-XXL	total
2.1,2.2	X-ray hard-band detected	326	574	1016	206	2122
2.3	Association with optical / IR	315	559	1016	206	2096
2.4	Redshift information:	180/131/11/4	227/322/15/10	491/519/0/6	174/0/32/0	1072/986/58/20
	spec-z/photo-z/no-z/removed stars					
2.5	X-ray spectral extraction and data analysis:	321/1	564/-	1010/-	206/-	2101/1
	successful/failed extraction					
2.6	galaxies removed (based on $L_{2-10\text{keV}} < 10^{42}\text{erg/s}$)	20	11	14	1	46
2.7	Objects used for LF analysis	302	553	996	205	2056

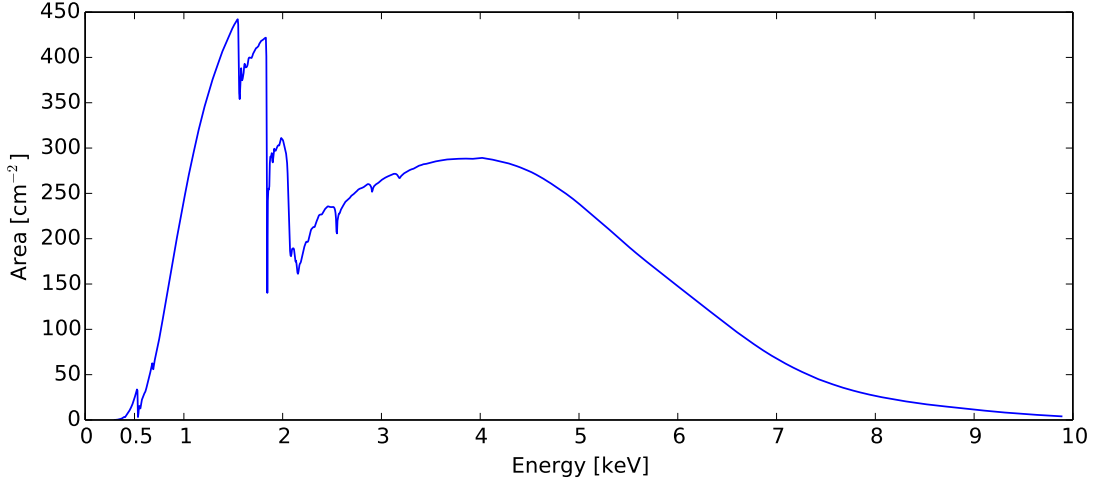


Figure 4.2.1: Typical response function (ARF) of the Chandra/ACIS detector. The area of the detector sensitive to a given energy (due to both mirror and detector effects) is plotted.

4.2 Detection of X-ray point sources

4.2.1 A detailed description for the CDFS

The CDFS survey region is, with 4 Ms total exposure time, the deepest observation performed by any X-ray telescope. The field of 465 arcmin² size was chosen for this purpose at high galactic latitude to minimise galactic absorption. The CDFS survey consists of 51 separate observations (“pointings”) which are aligned and merged to give the deep exposure necessary for detecting faint sources. During the data reduction of the individual observations other effects such as hot pixels, cosmic ray afterglows and times of anomalously high backgrounds (e.g. due to solar activity) have to be excluded. This yields “level-2 event files”, which are in essence a table where each row consists of a detected count (X-ray photon event), with its position on the sky, its energy and time of detection. As a side note, this makes X-ray astronomy unique in that with the same data, variability analysis (e.g. light curves), imaging and spectroscopy can be performed, or any combination thereof.

Not all X-ray photons are detected. In fact, the response of X-ray telescopes is very inefficient, and highly energy-dependent (see Figure 4.2.1). The cause of this is mainly the difficulty of focusing X-rays, which is achieved with grazing incidence mirrors. The response is formally modelled by a linear approximation: The incoming photon flux is split into discrete energy channels, and a matrix describes the response at the energy channels of the detector. This approach encapsulates both the efficiency of the system as well as the cross-talk between energy channels. This matrix is often separated into a normalised redistribution matrix (response matrix file, RMF) and a vector that captures the sensitivity of each energy channel (ancillary response file, ARF). These are either obtained by simulation or by calibration from on-board calibration sources. For Chandra,

a ray tracing software (MARX) is available that can simulate the response.

Also, the response of X-ray telescopes is dependent on the position of the source relative to the telescope axis. The focusing is best on-axis (mirror centre), where the point spread function (PSF) is smallest (albeit not necessarily symmetric). The size of the PSF increases towards the edges, while the efficiency of the mirror to focus X-rays decreases (vignetting). Furthermore, the imperfect alignment and rotation of the observations cause non-homogeneous exposure of the survey region. Additionally, fabrication limitations require small gaps between the charge-coupled devices (CCDs). All of these effects makes homogeneous detection across the inhomogeneously exposed region complicated.

The method for detecting X-ray sources used here follows Nandra et al. (2005) and was applied by Cyprian Rangel using the CIAO data analysis software (version 4.2) and custom software. Images and exposure maps in four energy bands (0.5–7, 0.5–2, 2–7 and 4–7 keV) were computed by aligning and merging the individual observations. First, a candidate source list was created using a wavelet-based detection algorithm (`wavdetect`) with a low significance threshold (10^{-4}). This produces a large number of candidate positions, many of which are likely spurious detections. In the next step, a statistical criterion was used to reduce this seed catalogue to secure detections. Source and background counts were extracted on the found positions using the Chandra point spread function tables of Laird et al. (2009). For each source, a source region was constructed with its size corresponding to 70% encircled energy fraction (EEF) of the point spread function (PSF). The background region is an annulus around the detected position with an inner radius between 1.5 times the 90% EEF of the PSF and a width of 100 pixels of 0.5". For determination of the background, other candidate sources are masked from the background extraction region. For each candidate source position, the Poisson probability that the observed counts are a background fluctuation is computed. The source is accepted if this probability is below 4×10^{-6} (Nandra et al. 2005). This work uses hard band (2 – 7 keV) selected sources. This is because this band is least biased to obscuration effects and therefore obscured AGN are detected. Additionally, this avoids non-AGN processes from galaxies (e.g. X-ray binaries and hot gas in star-bursting, but passive galaxies) which contribute to the X-ray flux in the soft band and thus contaminate the sample. This procedure yields a final catalogue of 569 sources for the CDFS.

This recipe for detection has the benefit that the probability of detecting a source (sensitivity) across the survey region can be estimated analytically. This is done by first computing the detection sensitivity at each position, which is defined as the number of counts necessary for a detection. This requires a continuous background count map, which is created by masking the detected sources from the image and applying smoothing and extrapolation¹. Next, the limiting flux in the hard band corresponding to the number of counts is calculated using the exposure maps in the 2 – 7 keV band following the methods described in Georgakakis et al. (2008). To compute the flux, a model has to be assumed for the spectrum of the source. Representing the “average” AGN spectrum, a powerlaw

¹Specifically, the areas surrounding each detected source is assumed to have no background information. The background in these areas is inferred from the neighbourhood.

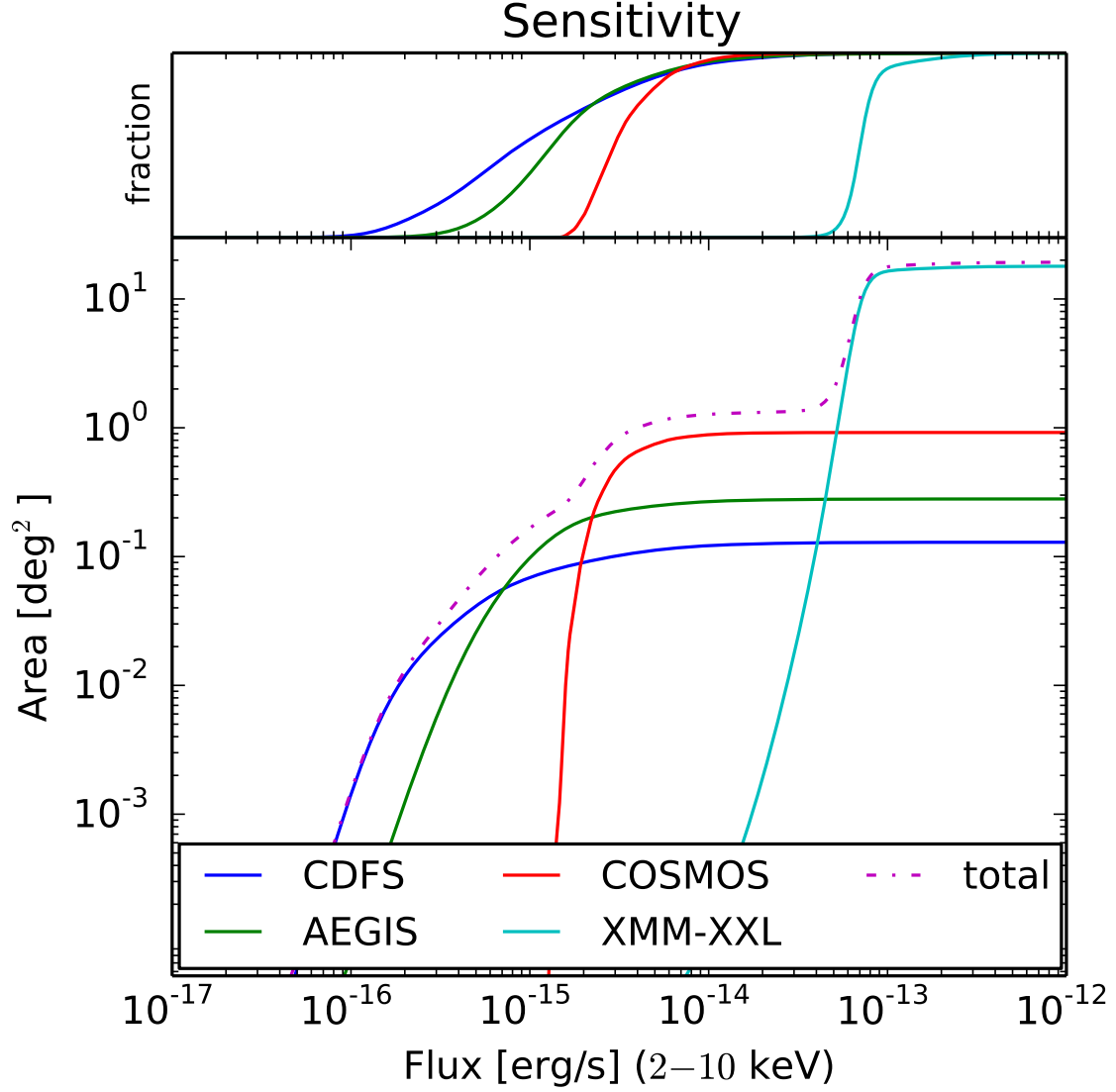


Figure 4.2.2: Area curve. The *bottom plot* shows all X-ray sensitivity curves on a logarithmic scale for the sum of all fields (dotted magenta line) and the individual fields. The *top plot* shows the sensitivity curves for the individual fields on a linear scale and normalised to their respective maximum area (see Table 4.1), to indicate the flux limit for detection. The XMM-XXL curve is limited in flux to 7×10^{-14} erg/s, but incorporates the uncertainty of measuring a higher flux due to Poisson variance. This introduces a smooth transition.

model with slope $\Gamma = 1.4$ is used. Finally, this map of limiting fluxes is collapsed into an single curve, such that for every 2 – 10 keV flux the area capable of detecting a source with this flux is given (the area curve, or sensitivity curve). In this computation, for each flux the area of each pixel is multiplied by the Poisson detection probability, consistent with the detection criterion mentioned in the previous paragraph, and summed over the survey area. The area curve of the CDFS field is shown in Figure 4.2.2.

4.2.2 Reduction of the other survey fields

I have begun by describing the data handling in the CDFS, because it is the field used in Chapter 3, and because the handling of the data in the other surveys is very similar. For the other three fields I now describe the differences. For the AEGIS-XD (Nandra et al., submitted) the data handling is the same as described above. For the C-COSMOS survey I use the 2 – 10 keV selected X-ray source catalogue presented by Elvis et al. (2009), which uses similar techniques. Only the XMM-XXL survey was treated substantially differently, and is described below. The number of sources detected in all four fields of choice, more than 2000 in total, is broken down by field and data reduction step in Table 4.2. The corresponding sensitivity curves are shown in Figure 4.2.2.

The XMM-XXL survey consists of pointings with 10 ks exposure, and covers 50 deg^2 in total, split into two fields of equal area. Here, data was used only from the equatorial sub-region due to the availability of spectroscopic redshifts there. This survey was included because it covers a large area, and thus samples the bright and rare AGN (as opposed to a detailed view of faint AGN in the deep Chandra fields). This is necessary for a complete view of the AGN population, specifically for constraining the bright end of the luminosity distribution (Chapter 6 and 7). Figure 4.2.3 illustrates the relative sizes of the used survey fields. The handling of the XMM-XXL survey is similar in spirit to the procedure described above for the Chandra surveys, but different due to technical reasons. The data reduction, source detection and sensitivity map construction follow the methods described in Georgakakis & Nandra (2011), and will be published in Liu et al. (in prep.). Here, the most important steps are outlined.

The XMM observations were reduced by Antonis Georgakakis using the Science Analysis System (SAS) version 12. The first step was to produce event files from the Observation Data Files (ODF) using the EPCHAIN and EMCHAIN tasks of SAS for the EPIC PN and MOS detectors respectively. Pixels along the edges of the CCDs of the PN and MOS detectors are removed because their inclusion often results in spurious detections. Flaring background periods are identified and excluded using a methodology similar to that described in Nandra et al. (2007a). Images and exposure maps in celestial coordinates with pixel size of 4.35 arcsec are constructed in 5 energy bands, 0.5 – 8, 0.5 – 2, 2 – 8, 5 – 8 and 7.5 – 12 keV. All overlapping EPIC images are merged prior to source detection to increase the sensitivity to point sources. The detection algorithm is applied independently to each of the 5 spectral bands defined above.

The source detection methodology is similar to that described in Laird et al. (2009) in the case of Chandra data. Source candidates are identified using the wavelet-based EWAVELET

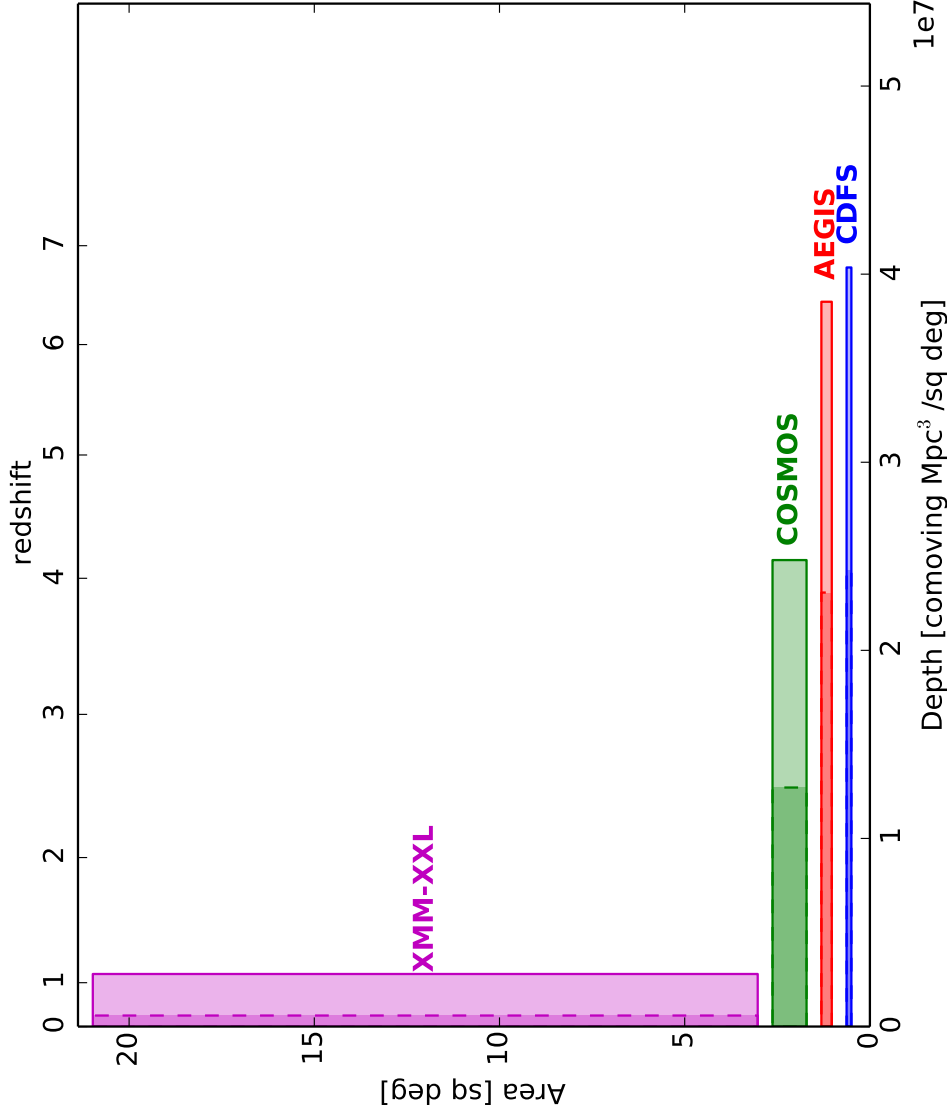


Figure 4.2.3: Volume captured by the observation fields. For this illustration, the area curve of each field (see Figure 4.2.2) was used to compute the maximum redshift at which at least 0.01 square degrees of the field can detect a $L = 10^{44.5}$ erg/s AGN with an obscuration column of $N_H = 10^{23} \text{ cm}^{-2}$ (a typical AGN). This redshift is converted into an integrated comoving volume per square degree. The area of the rectangles is thus proportional to the total volume sampled in each survey. The darker, dashed-encircled area depicts the same calculation but for $L = 10^{44}$ erg/s. The comparison highlights how the “shallow” XMM-XXL survey is useful for capturing the bright AGN population at $z < 1$, but is insensitive to faint AGN.

source detection task of SAS at a low threshold of 4σ above the background, where σ is the RMS of the background counts. For each candidate source the Poisson probability of a random background fluctuation was estimated. This step involved the extraction of the total counts at the position of the source and the determination of the local background value. To match the asymmetric PSF of XMM, especially off-axis, elliptical apertures were used from the XMM EPIC PSF parametrisation of Georgakakis & Nandra (2011). The count extraction region was obtained by scaling the elliptical apertures to contain 70 per cent of the PSF EEF. The total counts at a candidate source position, T , is the sum of the extracted counts from individual EPIC cameras. For each source the local background was estimated by first masking out all detections within 4 arcmin of the source position using an elliptical aperture that corresponds to the 80 per cent EEF ellipse. The counts from individual EPIC cameras were then extracted using elliptical annuli centred on the source with inner and outer semi-major axes of 5 and 15 pixels (0.36 and 1.09 arcmin) respectively, while keeping the same shape as the elliptical aperture in terms of rotation and ellipticity. The mean local background, B , was estimated by summing up the background counts from individual EPIC cameras after scaling them down to the area of the source count extraction region. The Poisson probability $P(T, B)$ that the extracted counts at the source position, T , are a random fluctuation of the background was calculated. Those sources with a low false detection probability $P(T, B) < 4 \times 10^{-6}$ are considered as detections. The above methodology is optimised for the detection of point sources. The final catalogue however, includes extended X-ray sources associated with hot gas from galaxy clusters or groups. Also, the extended X-ray emission regions of bright clusters are often split into multiple spurious detections by the source detection pipeline. The EMLDETECT task of SAS is used to identify extended sources (i.e. groups or clusters) and spurious detections. Point sources for which EMLDETECT failed to determine a reliable fit, are considered spurious and excluded from further analysis. The EPOSCORR task of SAS is used to correct for systematic errors in the astrometric positions of X-ray sources by cross-correlating with optical sources in the SDSS-DR8 catalogue (Aihara et al. 2011) with magnitudes $r < 22$ mag.

The flux of each source in different spectral bands is estimated by assuming a power-law X-ray spectrum with $\Gamma = 1.4$, i.e. similar to the XRB, absorbed by the appropriate Galactic hydrogen column density. The latter is derived from the HI map of Kalberla et al. (2005) using the right ascension and declination of the aimpoint of each XMM observation and the NH task of FTOOLS. The energy to flux conversion factors are chosen to transform counts from the 0.5-2, 0.5-8, 2-8, 5-8 and 7.5-12 keV bands to fluxes in the 0.5-2, 0.5-10, 2-10, 5-10 and 7.5-12 keV bands respectively.

From the XMM-XXL point source catalogue selected in the 2 – 8 keV band, I use a subsample with a total of 206 sources (see Table 4.2). This sample was chosen by maximising the completeness of optical identification and spectroscopic redshift determination (see next sections) through applying a bright flux cut, $f_X(2 - 10 \text{ keV}) > 7 \times 10^{-14} \text{ erg s}^{-1} \text{ cm}^{-2}$.

The computation of area curves follows the procedure described before for the CDFS. Here, additionally the flux cut has to be taken into account by modifying the number of counts necessary for detection when computing the sensitivity map. At each survey position, the probability for detecting a source with a given flux and measuring a flux above

the flux limit was computed. Due to the Poissonian nature of photon counts (Eddington bias), the latter condition still introduces a smooth transition in the area curve, shown in Figure 4.2.2. Specifically, it is possible to have a fainter object in the survey field, but estimate a flux above the flux limit due to count fluctuations, and thus including it in the sample. The area curve reflects this properly.

4.3 Extraction of spectra

Now that the point sources have been detected, extraction of the X-ray spectra is the next step. The `ACIS EXTRACT` (AE) software package (Broos et al. 2010) was used to extract spectra for each source. Initially, each source and each pointing is dealt with separately. AE simulates the PSFs at each source position using MARX. Regions enclosing 90% PSF EEf at 1.5keV were used to extract source spectra. The background regions are constructed around the sources such that they contain at least 100 counts, with other sources masked out. AE also constructs local RMFs and ARFs using ray-tracing. As a final step, AE merges the extracted spectra so that each source has a single source spectrum, a single local background spectrum, and a single ARF and RMF. For two sources detected at the very edge of the CDFS exposure region, the spectral extraction failed. These sources were not used, leaving 567 sources for the CDFS.

In the XMM-XXL field, a similar approach was used, but the extraction regions were chosen manually (see Liu et al., in prep.), and the RMF/ARFs are provided by the calibration within the SAS software package.

The next step is the spectral analysis of the X-ray spectra. However, before doing scientific analysis, the contribution of the background to the spectrum has to be considered (Section 4.4 below) as well as the estimation of redshifts (Section 4.7) via the association to optical/infrared counterparts (Section 4.5).

4.4 Background model

For consistent analysis using Poisson statistics, a model has to be compared to the observed counts. The background contribution can not be subtracted away because unlike in Gaussian distributions, the subtraction of two Poisson distributions does not yield an analytic distribution. It is common practise to use per-bin background estimates. However, this yields unstable results in bins with few counts. I thus choose to model the background in a parametric way using a Gaussian mixture model in the Chandra fields. This may not be a physical model but it provides a good approximation to the background which is very similar to the on-orbit background measurements (Baganoff 1999).

In the XMM-XXL field, this background model is not appropriate. There, a model specific to the instruments of XMM-Newton was used, which was presented in the doctoral thesis of R. Sturm (Sturm 2012) in a different context. This model is also based on a collection of Gaussian lines. However, the continuum background is modeled with spline

functions.

4.4.1 Background model definition

The background model in the Chandra fields consists of a number of instrumental emission lines represented by Gaussian components at mean energies of 1.486 (Al-K α), 1.739 (Si K α), 2.142 (Au M α, β), 7.478 (Ni K α), 9.713 (Au L α ; all in keV, Thompson et al. 2001). The centres of these lines are allowed to vary within 0.1 keV. A feature at ~ 8.3 keV (possibly *Ni/Au* lines) is described by three additional Gaussians at 8.012, 8.265, 8.494 keV. The overall continuum shape is modelled with a flat continuum level. Additionally, to model the rise in the soft energies two Gaussians (called “softend” and “softsoftend”) are used of widths ~ 0.5 keV and ~ 2 keV at centred at ~ 0 keV. All background model parameters (means, widths, heights) are then fitted to the extracted background spectrum of each source. This Gaussian mixture model, “Model 1”, can be written as

$$M_1(E) = C \times \left(1 + \sum_i A_i \times \exp \left\{ -\frac{1}{2} \left(\frac{E - E_i}{FWHM_i} \right)^2 \right\} \right), \quad (4.4.1)$$

with a constant base continuum C and Gaussian components of relative importance to the continuum A_i , central energy E_i and full width half maximum parameter $FWHM_i$. All centres are allowed to vary within 0.1 keV. The parameters start from reasonable guesses and are optimised as long as the fit statistic (C-Stat) improves.

Figure 4.4.1 shows the comparison between data and model for the background spectrum of source 318 in the CDFS catalogue. The background spectrum has a total of 3380 counts and was chosen because it constitutes an intermediate case between the most high-count spectra with many peculiarities and low-count spectra with almost no visible features. The final parameter values after fitting are shown in Table 4.3 (middle column).

I now try to demonstrate that the background description is a good model. To this end, a goodness of fit (GoF) methodology for X-ray spectra is presented. I use Q-Q plots for model discovery and the AIC model comparison method to test for the significance of model improvements, although any of the model comparison methods introduced in Section 5.2.2 could be chosen (see references and discussion there). I demonstrate the method using a background source spectrum and the best fit, comparing it to a simplified model. In my analysis, every spectrum is fitted individually to accommodate the diversity of background spectra.

4.4.2 Goodness of Fit and Model discovery

For comparison, I present Model 2 which has several Gaussians disabled, namely the one centred at 2.1 keV (line 3) and the three between 8 – 9 keV (line 5-7). The intent is to compare methods evaluating whether Model 2 is a good model, where it deviates from the data, and whether the deviation is significant.

CDFS Source ID	318	179
C	7.278×10^{-5}	23.28×10^{-5}
line1.ampl	413	511
line1.fwhm	0.0329	0.0162
line1.pos	1.475	1.49
line2.ampl	0	74.7
line2.fwhm	0.1	0.002
line2.pos	1.84	1.82
line3.ampl	420	581
line3.fwhm	0.1	0.0827
line3.pos	2.16	2.16
line4.ampl	3769	1159
line4.fwhm	0.0257	0.0768
line4.pos	7.48	7.49
line5.ampl	1446	42.9
line5.fwhm	0.002	0.102
line5.pos	8.10	8.07
line6.ampl	97.1	280
line6.fwhm	0.4	0.0793
line6.pos	8.37	8.25
line7.ampl	111.5	6656
line7.fwhm	0.002	0.002
line7.pos	8.47	8.47
line8.ampl	1993	2225
line8.fwhm	0.1	0.0886
line8.pos	9.72	9.71
softend.ampl	21092	21824
softend.fwhm	2.0	2.6
softend.pos	1.0	1.0
softsoftend.ampl	72642	78711
softsoftend.fwhm	0.325	0.638
softsoftend.pos	0.167	0.243

Table 4.3: Background model parameters for two sources. For the left source, the model is shown in Figure 4.4.1. Positions and FWHM are in keV. Amplitudes are unitless, except for the normalisation C , which is in cts/s/keV/cm^{-2} .

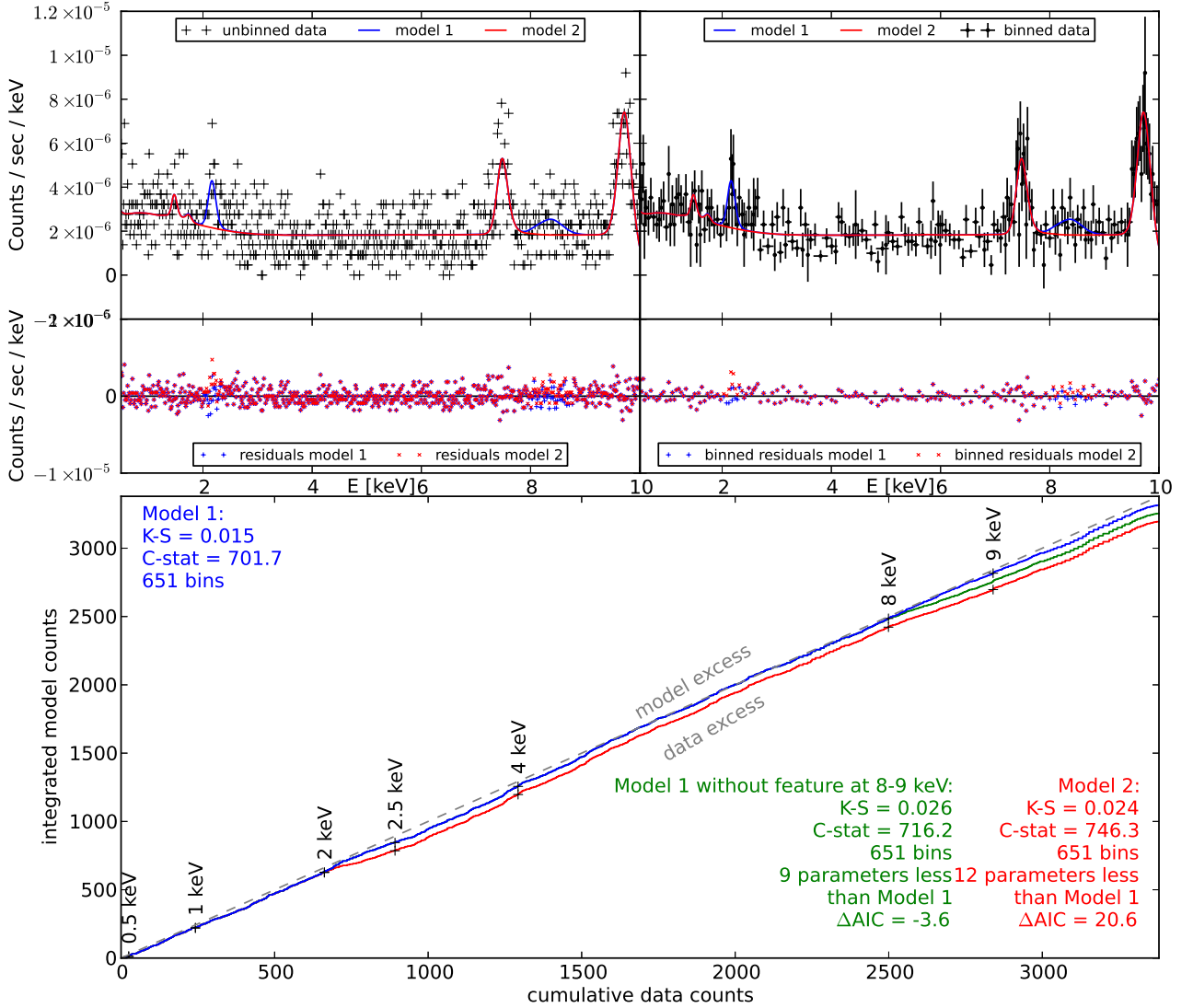


Figure 4.4.1: Comparison of the background data from Source 318 with background models. The best model is shown in blue, while the red model has several features removed (see text in Section 4.4). In the top two panels, the usual count spectrum is shown with residuals (left unbinned, right binned to at least 20 counts per bin). The large, bottom panel presents the corresponding Q-Q (quantile-quantile) plot. For each energy E , the model counts predicted and the counts observed below E are recorded on the plot. The grey dashed line is where data and model would perfectly agree. Model 1 (blue solid top line) follows this line very closely, and thus can be considered a good model. Model 2 (red solid bottom line) deviates from the grey dashed line at 2 keV, indicating that a feature in the data may be missing in the model. The shape and size of the deviation also indicates the shape of the needed feature. The significance of the feature can be tested using model selection. Here, the AIC shows that the feature at 2 – 2.5 keV is justified ($\Delta AIC > 0$), but adding the another feature at 8 – 9 keV is not (shown in green, see text in Section 4.4 for details).

The classic method is to plot the data, model and the residuals. This is shown in the upper left panel of Figure 4.4.1 for the un-binned data and model, with the residuals below. The upper right panel shows the same, but with adaptive binning requiring 20 counts in each bin. The feature at 2.1 keV is visible immediately, while the feature at 8 – 9 keV is less striking.

I introduce an alternative method of analysing the quality of a model: the **Q–Q plot**, shown in the large lower panel. For each energy E , the model counts predicted and the counts observed below E are recorded on the plot. Here the counts are shown, while statisticians typically use quantiles, i.e. the fraction of observations that lie below a quantity. This does not influence the main point, namely the shape of the curve. The grey dashed line is where data and model would perfectly agree. A steeper curve means the model predicts more counts than observed, while a shallower curve indicates an excess of observed counts.

An example of the Q-Q plot is shown in the lower (large) panel of Figure 4.4.1. The gray dashed line indicates the line where the model counts agree perfectly with the observed counts. Model 1 (blue solid top line) follows this ideal line very closely, and thus can be considered a good model. When removing a feature, as in Model 2 (red solid bottom line), deviations from the grey dashed line can be seen, indicating that a feature in the data (here at 2 keV) is not modelled. Above 2.5 keV, the line is parallel to Model 1, indicating no further difference. This means the feature is confined to this energy range. One can also see that the difference required to bring the lines into agreement looks like the cumulative distribution of a Gaussian (a S-shape) rather than e.g. a straight line for a flat distribution. Another, but more subtle, deviation is visible between 8 – 9 keV. This is highlighted using the green solid middle line which does model the 2 keV feature.

Having found a good model (Model 1), and slightly worse, simpler models, the next step is to test whether the improvement is significant. For instance, it seems doubtful that the small feature at 8 – 9 keV justifies modelling with 3 Gaussian components (9 parameters). For this, the AIC is employed, which punishes the improvement in likelihood (C-stat) by the number of parameters. This technique is presented in Section 5.2.2 in more detail. As $AIC(\text{Model 1}) - AIC(\text{Model 2}) = 20.6 > 0$, the worsening of the fit is strong enough to justify the feature at 2 keV (red text in Figure 4.4.1). But if instead the feature at 8 – 9 keV is removed, the AIC decreases due to the simplification of the model (green text). Thus, in this source, the 8 – 9 keV feature can be ignored. However, in sources with more counts, it is required (see Figure 4.4.2 for one example). Based on this evaluation (across all background spectra), I adopt “Model 1” for the Chandra surveys, even for individual spectra where the model not justified. This is because it always captures the necessary features, regardless of the number of counts in the background spectrum.

4.4.3 Using the background model

For the subsequent X-ray spectral analysis of the source spectrum, the background spectral model specified above (Model 1) is added on top of the source spectral model, scaled by the relative size of the background and source extraction regions. The inferred parameters

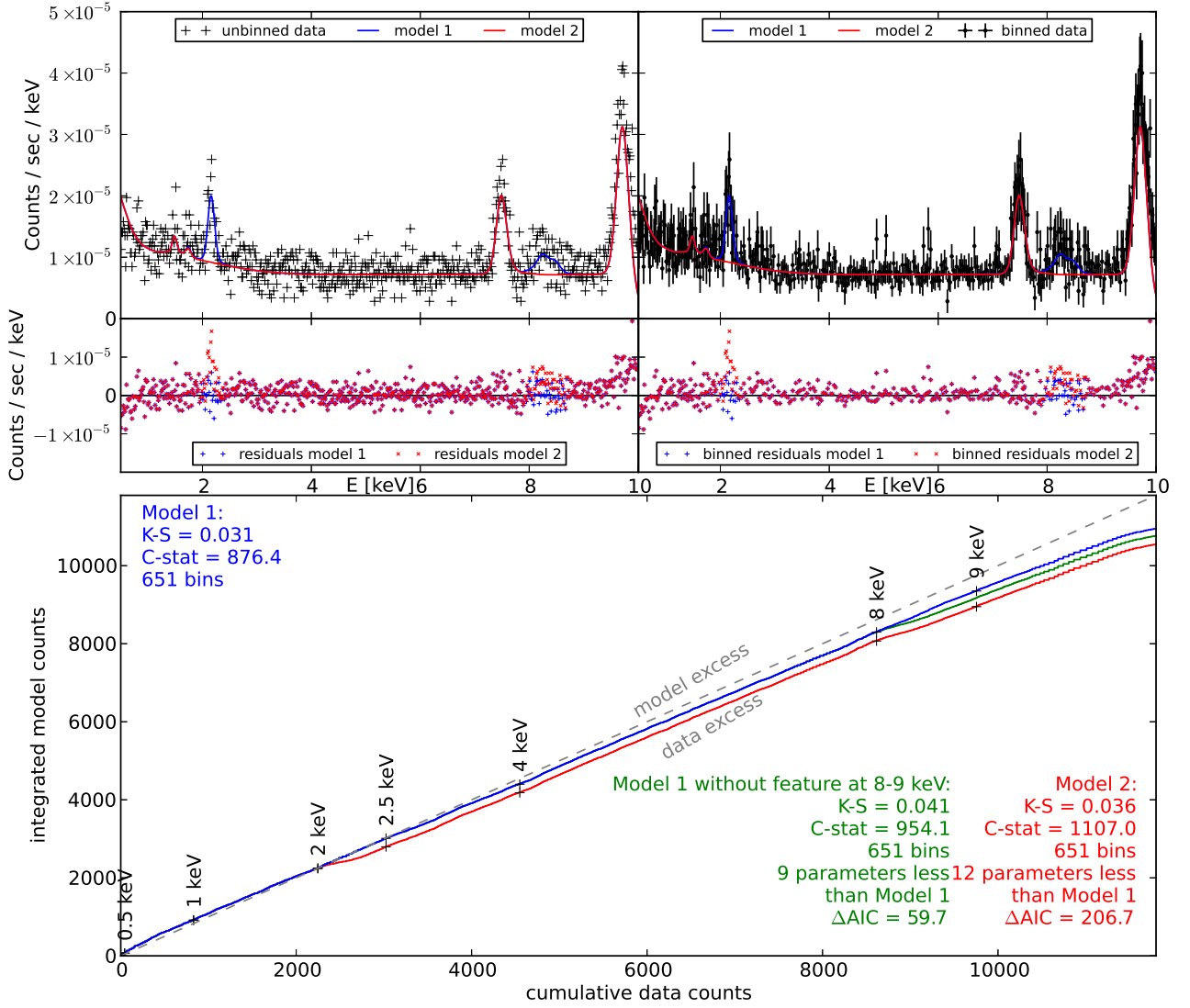


Figure 4.4.2: Same as Figure 4.4.1, but for Source 179 (11802 counts). Unlike Figure 4.4.1, the feature between 8 – 9 keV is required ($\Delta AIC > 0$). In the lower panel, there is a mild, continuous deviation from the grey dashed line indicating that a mild increase in the higher energy counts. This hints that the model could potentially be improved further, e.g. by applying a slope to the continuum. Because this spectrum has the highest number of counts, is the only one exhibiting this deviation, and source spectra typically have one order of magnitude fewer counts than background spectra, I did not attempt further complications of the model.

for the background spectral model are kept fixed (although they are different from source to source). In general, there may be cases where simultaneous analysis of background and source model parameters provides better results. However, because the larger background region captures many more background photons than the source region, the background model is well constrained by the background data alone.

4.5 Association with optical/infrared counterparts

For a meaningful analysis of the X-ray spectra, redshift information is needed. Due to the absence of clear lines in the X-ray spectrum of AGN and the complicated form of the response, the redshift can generally not be obtained from the X-ray spectrum directly. It is thus necessary to associate the X-ray position of each source with optical/infrared counterparts and estimate redshifts using photometric methods or optical/infrared spectroscopy (see Section 4.7).

For the association of AGN, infrared wavelengths provide the best matching capabilities, because the dusty torus of AGN is often a strong black-body emitter in these wavelengths which stands out due to the slightly lower number of other sources emitting in this band as compared to other wavelengths and the high resolution of deep surveys by infrared telescopes, making the matching very accurate.

In the AEGIS-XD, COSMOS and XMM-XXL fields, I used published associations. There, the Likelihood Ratio method of Sutherland & Saunders (1992) was used for the identification of the X-ray sources with optical or infrared counterparts. In this method, the distance between detected positions between the two catalogues is considered. Additionally, the likelihood is weighted by the ratio $\bar{q}(m)/\bar{n}(m)$, where m is the magnitude in the considered wavelength band, \bar{q} is the magnitude distribution of the target population (e.g. AGN, known from previous surveys) and \bar{n} is the magnitude distribution of the background sources (all others, e.g. stars, passive galaxies). Specific details on the association of X-ray sources with optical/infrared counterparts are presented by Nandra et al. (submitted). They used the multi-waveband photometric catalogue provided by the Rainbow Cosmological Surveys Database (Pérez-González et al. 2008; Barro et al. 2011a,b). The counterparts of C-COSMOS X-ray sources are taken from Civano et al. (2012). For the identification they used the I -band selected optical sample of Capak et al. (2007), the K -band photometry of McCracken et al. (2010) and the IRAC-3.6 μm catalogue of Sanders et al. (2007). X-ray sources in the XMM-XXL survey were matched to the SDSS-DR8 photometric catalogue Aihara et al. (2011) following the methods described in Georgakakis & Nandra (2011).

For the CDFS, a different association method was used. I developed a Bayesian version of the Likelihood Ratio method based on the probabilistic formalism of Budavári & Szalay (2008), which is presented in the next section. This method has the advantage that multiple catalogues can be taken into account easily. However, results in the CDFS are very similar to those obtained with the Likelihood Ratio method. This method was applied by Li-Ting Hsu in the CDFS region, and presented in Hsu et al. (2014) as a catalogue of the

counterparts to CDFS X-ray sources. The photometric catalogues used include the CANDELS multi-wavelength catalog of Guo et al. (2013), the MUSYC catalogue presented by Cardamone et al. (2010) and the TENIS near-infrared selected source catalogue described by Hsieh et al. (2012).

In the following section I present the new Bayesian method in detail. Finally, I have found that for some fraction of objects it is actually possible to obtain the redshift from the X-ray spectrum directly. This is discussed in Section 4.7.2.

4.6 Bayesian association method (Nway)

Lets consider the problem of finding associations across catalogues. I strongly rely on the methodology developed by Budavári & Szalay (2008) to compute the probability of an association between n catalogues based on distances (described in Section 4.6.1). However, I do not require an object to have a counterpart in each of the catalogues. Especially for faint sources, it is often the case that the counterpart is not detected. In Section 4.6.1, I describe how unlikely associations are removed, revealing the most likely counterpart, or multiple options in ambiguous cases.

In Section 4.6.2, I amend this approach by also considering magnitude information. Members of a certain class of objects will have a different magnitude/colour distribution than the non-members of this class. Thus, the results can be informed about more probable associations based on previous knowledge. Finally, in 4.6.4, I describe the NWAY implementation which implements the methodology in a generic, easy-to-use fashion.

4.6.1 Distance-based matching

In this work, I consider the problem of finding counterparts to a primary catalogue ($i = 1$), namely the X-ray source position catalogue. Let each N_i denote the number of entries for the catalogues used, and $\nu_i = N_i/\Omega_i$ denote their source surface density on the sky.

If a counterpart is required to exist in each of the k catalogues, there are $\prod_{i=1}^k N_i$ possible associations. If we assume that a counterpart might be missing in each of the matching catalogues, there are $N_1 \cdot \prod_{i=2}^k (N_i + 1)$ possible associations. This minor modification, negligible for $N_i \gg 1$, is ignored in the following for simplicity.

If each catalogues covers the same area with some respective, homogeneous source density ν_i , the probability of a chance alignment on the sky of physically unrelated objects can then be written (Budavári & Szalay 2008, eq. 25) as

$$P(H) = N_1 / \prod_{i=1}^k N_i = 1 / \prod_{i=2}^k N_i = 1 / \prod_{i=2}^k \nu_i \Omega_i. \quad (4.6.1)$$

Thus $P(H)$ is the prior probability of an association. The posterior should strongly exceed this prior probability, to avoid false positives.

To account for non-uniform coverage, $P(H)$ is generalised so that the “prior completeness factor” c , which gives the expected fraction of sources with reliable counterpart (due to

only partial coverage of the matching catalogues $\Omega_{i>1} \neq \Omega_1$, depth of the catalogues and/or systematic errors in the coordinates). Our prior can thus be written as

$$P(H) = c / \prod_{i=2}^k \nu_i \Omega_1. \quad (4.6.2)$$

Bayes' theorem connects the prior probability $P(H)$ to the posterior probability $P(H|D)$, by incorporating information gained from the observation data D via

$$P(H|D) \propto P(H) \times P(D|H). \quad (4.6.3)$$

Then, comparing two different hypotheses (H and \bar{H}), we can write how probable one is compared to the other as

$$\frac{P(H|D)}{P(\bar{H}|D)} \propto \frac{P(H)}{P(\bar{H})} \times \frac{P(D|H)}{P(D|\bar{H})} \quad (4.6.4)$$

$$= \frac{P(H)}{P(\bar{H})} \times B \quad (4.6.5)$$

where the Bayes factor B indicates the strength of hypothesis H based on the observations. In the case considered here, for each association, the model “chance alignment” is compared with the model “real association”. The relevant Bayes factor, developed in (Budavári & Szalay 2008, eq. 18), is dependent on the angular distance ϕ_{ij} between the source positions in catalogues i and j :

$$\begin{aligned} B &= \frac{P(\text{"real association"}|D)}{P(\text{"chance alignment"}|D)} \\ &= 2^{n-1} \frac{\prod \sigma_i^{-1}}{\sum \sigma_i^{-1}} \exp \left\{ -\frac{\sum_{i=1}^j \phi_{ij} \sigma_j^{-1} \sigma_i^{-1}}{\sum \sigma_i^{-1}} \right\} \end{aligned} \quad (4.6.6)$$

Here, σ_i denotes the positional uncertainty in each catalogue.

For each association the posterior of the hypothesis “real association” is then

$$P(\text{"real association"}|D) = \left[1 + \frac{1 - P(H)}{B \cdot P(H)} \right]^{-1} \quad (4.6.7)$$

For each combinatorically possible association across the catalogues, Equations 4.6.7 and 4.6.6 describe the probability that this association is real (i.e. not by chance). The task now is to begin with the primary catalogue (the X-ray positions) and find for each X-ray detection one or more realistic associations to consider.

First, I write the probability that the X-ray source has any association at all, by summing the probability of its possible counterparts:

$$P(\text{"any real association"}|D) = \left[1 + \frac{1 - P(H)}{\sum_k B_k \times P(H)} \right]^{-1} \quad (4.6.8)$$

This is possible because the prior remains the same across all associations and because associations are mutually exclusive. Based on this posterior we can reject or accept the hypothesis that a real association exists for each object in the primary catalogue. If we accept the hypothesis that this source has a counterpart, the probability for any specific association k is

$$P(k|D, \text{"any real association"}) = B_k / \sum_k B_k. \quad (4.6.9)$$

Here, I have used that a priori all associations are equally likely. This posterior can be used to reject unlikely counterparts. A “secure” counterpart could be defined by the requirement $P(\text{"any real association"}|D) > 99\%$ and $P(k|D, \text{"any real association"}) > 95\%$, for example.

One subtlety of Equation 4.6.8 is that it only considers the associations enumerated, i.e. the combinations of all detected objects. However, very faint counterparts may have not been detected with the current exposure depth. The possibility that another, undetected counterpart is the correct one is thus not included in Equation 4.6.8 nor Equation 4.6.9. Low probabilities in Equation 4.6.7 even for the most probable association may indicate that no suitable counterpart has been found yet.

4.6.2 Magnitudes

Astronomical objects of various classes often show distinct color and magnitude distributions. Because most X-ray point-sources in deep images are AGN, and the fact that AGN show a different optical magnitude distribution than other objects such as stars, this information can be exploited. Previous works (Brusa et al. 2005, 2007) have modified the likelihood ratio coming from the angular distance $f(r)$ information (likelihood ratio method, Sutherland & Saunders 1992) by a factor:

$$LR = \frac{q(m)}{n(m)} \times f(r) \quad (4.6.10)$$

Here, $q(m)$ and $n(m)$ are associated with the magnitude distributions of source (e.g. AGN) and background objects (e.g. stars, passive galaxies) respectively, but additionally contain sky density contributions.

This idea can be put on solid footing within the Bayesian framework. Here, two Bayes factor are combined, by simply considering two independent observations, namely one for the positions, D_ϕ , and one for the magnitudes D_m . The Bayes factor thus becomes

$$B' = \frac{P(D_\phi|\text{"real association"})}{P(D_\phi|\text{"chance alignment"})} \times \frac{P(D_m|\text{"real association"})}{P(D_m|\text{"chance alignment"})} \quad (4.6.11)$$

$$= B \times \frac{\bar{q}(m)}{\bar{n}(m)}, \quad (4.6.12)$$

with $\bar{q}(m)$ and $\bar{n}(m)$ being the probability that a X-ray source or a background object has magnitude m respectively.

For completeness, I mention the fully generalised case. This is attained when an arbitrary number of photometry bands are considered, each consisting of a magnitude measurement m and measurement uncertainty σ_m :

$$B' = B \times \prod \frac{\int_m \bar{q}(m) p(m|D_m) dm}{\int_m \bar{n}(m) p(m|D_m) dm} \quad (4.6.13)$$

Here, $p(m|D_m)$ would refer to a Gaussian error distribution with mean m and standard deviation σ_m . This is convolved with the distribution properties. Alternatively, $p(m|D_m)$ can also consider upper limits. The posterior formulae $P(\cdot|D)$ introduced above (Equations 4.6.8 and 4.6.9) remain the same, with B' replacing B .

4.6.3 Auto-calibration

The probability distributions $\bar{q}(m)$ and $\bar{n}(m)$ can be taken from other observations by computing the magnitude histograms of the target population (e.g. X-ray sources, AGN) and other objects selected in other wavebands.

Under certain approximations and assumptions, these histograms can also be computed during the catalogue matching procedure while also being used for the weighting. One could perform the distance-based matching procedure laid out above, and compute a magnitude histogram of the secure counterparts as an approximation for $\bar{q}(m)$ and a histogram of ruled out counterparts for $\bar{n}(m)$. While the weights $\bar{q}(m)/\bar{n}(m)$ may strongly influence the probabilities of the associations for a single object, the bulk of the associations will be dominated by distance-weighting. One may thus assume that the $\bar{q}(m)$ and $\bar{n}(m)$ are computed with and without applying the magnitude weighting are the same, which is true in practice. When differences are noticed, they will only strengthen $\bar{q}(m)$, and the procedure may be iterated.

4.6.4 Implementation

My implementation for matching n catalogues is a Python program called NWAY. The input catalogues have to be in FITS format. Information about the (shared) sky coverage has to be provided to the program as well. The program proceeds in four steps.

First, possible associations are found. It is unfeasible to consider all theoretical possibilities (complexity $O(\prod_{i=1}^k N_i)$), so the sky is split first to cluster nearby objects. For this, a hashing procedure puts each object into square bins. The bin width w is chosen so that an association of distance w is improbable, i.e. much larger than the largest positional error. An object with coordinates ϕ, θ is thus put into bin $(i, j) = (\lfloor \phi/w \rfloor, \lfloor \theta/w \rfloor)$, but also into bins $(i+1, j)$, $(i, j+1)$ and $(i+1, j+1)$ to avoid boundary effects. This is done for each catalogue separately. Then, in each bin, the Cartesian product across catalogues (every possible combination of sources) is computed. All associations are collected across the bins

and filtered to be unique. The hashing procedure adds very low effort $O(\sum_{i=1}^k N_i)$ while the Cartesian product is reduced drastically to $O(N_{bins} \cdot \prod_{i=1}^k \frac{N_i}{N_{bins}})$. All primary objects that have no associations past this step have $P(\text{"any real association"}|D) = 0$.

The second step is the computation of Bayes factors using the angular distances between counterparts. The prior is also evaluated from the size of the catalogue and the effective coverage, as well as the user-supplied prior incompleteness factor. The posterior for each association based on the distances only is calculated.

In the third step the magnitudes are considered, and the Bayes factors modified. An arbitrary number of magnitude columns in the input catalogues can be specified. It is possible to use external magnitude histograms (e.g. for sparse matching with few objects) as well as computing the histograms from the data itself (see Section 4.6.3). The breaks of the histogram bins are computed adaptively based on the empirical cumulative distribution found. Because the histogram bins are usually larger than the magnitude measurement uncertainty, the latter is currently not considered. The adaptive binning creates bin edges based on the number of objects, and is thus independent of the chosen scale (magnitudes, flux). Thus the method is not limited to magnitudes, but can be used for virtually any other known object property (colours, morphology, variability, etc.).

In the final step, associations are grouped by the object from the primary catalogue (here: the X-ray source). The posteriors $P(\text{"any real association"}|D)$ and $P(k|D, \text{"any real association"})$ computed. For the output catalogue a cut on the posterior probability (e.g. above 80%) is applied, and all associations with their posterior probability are written to the output fits catalogue file. In practice, for the case of the CDFS field, the automated procedure laid out above results in a single association for the majority of X-ray sources. For a small number of sources, no counterpart, or multiple possible counterparts are found (see Hsu et al. 2014, for more details).

4.7 Redshift estimation

The X-ray survey fields used in this work benefit from extensive spectroscopic campaigns that also specifically target X-ray sources. In the CDFS, the spectroscopic redshifts compiled by N. Hathi (private communication, see Hsu et al. 2014) are used. Spectroscopic redshift measurements of X-ray sources in the AEGIS field are extracted from the compilation presented in Nandra et al. (submitted) which included also the DEEP2 (Newman et al. 2012) and DEEP3 galaxy redshift surveys (Cooper et al. 2011, 2012) as well as observations carried out at the MMT using the Hectospec fibre spectrograph (Coil et al. 2009). Redshifts in the C-COSMOS are used from the compilation of Civano et al. (2012) which includes the public releases of the VIMOS/zCOSMOS bright project (Lilly et al. 2009) and the Magellan/IMACS observation campaigns (Trump et al. 2009). In the case of XMM-XXL, optical spectroscopy is from Stalin et al. (2010), the Baryon Oscillation Spectroscopic Survey (BOSS; Dawson et al. 2013; Bolton et al. 2012; Smee et al. 2013) as well as dedicated Sloan Digital Sky Survey III (SDSS; York et al. 2000; Gunn et al. 2006; SDSS-III: Eisenstein et al. 2011) ancillary science observations, which targeted specifically

X-ray sources in the equatorial XMM-XXL field (PI: A. Merloni, A. Georgakakis). Targets were selected to have $f_X(0.5 - 10 \text{ keV}) > 10^{-14} \text{ erg s}^{-1} \text{ cm}^{-2}$ and $17 < r < 22.5$, where r corresponds either the PSF magnitude in the case of optical unresolved sources (SDSS type=6) or the model magnitude for resolved sources (see Menzel et al., in prep.). At the flux limit $f_X(2 - 10 \text{ keV}) > 7 \times 10^{-14} \text{ erg s}^{-1} \text{ cm}^{-2}$, 84% (174/207) of the XMM-XXL sources have secure redshift measurements.

Overall, only 1072 of the 2096 X-ray sources have spectroscopic redshifts (see 4.2). The CDFS, AEGIS-XD and COSMOS fields have multiwavelength photometric observations that allow photometric redshift estimates for sources that lack spectroscopy. The photometric redshifts were computed by Li-Ting Hsu and Mara Salvato and are presented in Hsu et al. (2014) for the CDFS, Nandra et al. (submitted) for the AEGIS-XD field, and Salvato et al. (2011) for the COSMOS field. These employed methods developed in Salvato et al. (2009, 2011) to achieve photometric redshift accuracies for X-ray AGN comparable to galaxy samples. This photometric fitting method, using software from Ilbert et al. (2006, 2009), is capable of not just computing a most likely redshift point estimate, but also produces redshift probability distributions. For the study in Chapter 5, I used the photo-z probability distributions directly to incorporate the uncertainty of the redshift estimate. However, in addition to the uncertainty due to measurement errors, the method of photometric redshift has systematic errors due to incomplete template libraries, failures in automatically extracting correct fluxes at the edge of images, blending of sources causing the flux to be confused, etc. The most important systematic contribution is probably incorrect associations. In Section 4.7.1, I show how systematic effects can be taken into account to make further inference more robust. The outcoming photometric distributions were used for the analyses in Chapter 6.

A by-product of the photometric redshift determination is the characterisation of the Spectral Energy Distribution (SED) of X-ray sources. This information is used to identify candidate galactic stars in the X-ray sample and exclude them from the analysis. Following Salvato et al. (2011) if stellar templates provide an improved fit to the SED of a source, as measured by the reduced χ^2 i.e. $\chi_{\text{star}}^2 < \chi_{\text{gal}}^2/1.5$, and the source is point-like in the optical images, then it is considered to be a galactic star candidate. The number of removed stars is indicated in Table 4.2.

For the XMM-XXL field, multiwavelength photometry from the UV to the infrared which homogeneously covers the surveyed area is not available. This is essential for reliable photometric redshifts, especially in the case of bright X-ray samples like the XMM-XXL (e.g. Salvato et al. 2011). Therefore, for the X-ray sources without spectroscopic redshift measurement in that field (32/206) I chose not to determine photometric redshifts. These sources are still included in the analysis by assigning them a flat redshift prior (see below for details). Inspection of the images suggests no contaminating stars in the XMM-XXL bright sub-sample.

To summarise, the redshift determination falls in one of three cases for each X-ray source:

1. *No redshift information available* (58 sources): This is the case if no spectroscopic redshifts are available, and photometric redshift estimation is not possible due to

limited photometry. This case also applies when no secure association was found. Such a source is associated with a flat redshift distribution in the interval $z = 0.001 - 7$.

2. *Spectroscopic redshifts* (1072 sources): Wherever secure spectroscopic redshifts are available, they are used directly, i.e. I do not associate redshift uncertainty to them. Different surveys use different conventions to define the reliability of the redshift measurement. This work only considers spectroscopic redshifts from the top two quality classes of any study, which typically corresponds to a probability better than 95 per cent of being correct.
3. *Photometric redshifts* (986 sources): These are included in the analysis in the form of probability distribution functions. I also incorporate systematic uncertainties e.g. due to incorrect association to an optical counterpart (see Section 4.7.1).

Table 4.2 on page 35 shows a breakdown of how many sources fall into each category. The distribution of redshifts is plotted in Figure 4.7.1 for spectroscopic and photometric redshifts respectively. Figure 4.7.2 shows the redshift distribution in each field.

4.7.1 Robust photometric redshift probability distributions

On a high-level, the quality of redshift point estimates for a specific sample can be described by two quantities, which are estimated from a sub-sample where spectroscopic redshifts are available: the outlier fraction η and the scatter σ (Salvato et al. 2011). The outlier fraction describes the failure rate or accuracy of the redshifts, i.e. how frequently the redshift estimate is far off the mark:

$$\eta : \text{fraction where } \left| \frac{z_{\text{phot}} - z_{\text{spec}}}{1 + z_{\text{spec}}} \right| > 0.15 \quad (4.7.1)$$

The scatter width σ describes how precisely the redshifts are scattered around the true value. The sample can be characterised using the normalised median absolute deviation (NMAD, Hoaglin et al. 1983) as a robust estimator:

$$\sigma_{\text{NMAD}} = 1.48 \times \text{median} \left| \frac{z_{\text{phot}} - z_{\text{spec}}}{1 + z_{\text{spec}}} \right| \quad (4.7.2)$$

These two quantities are interpreted such that the redshift estimate z_{phot} is distributed along a normal distribution centred at the true value z_{spec} with standard deviation $\sigma_{\text{NMAD}} \cdot (1 + z_{\text{spec}})$. Additionally, a fraction η of the redshift estimates are drawn from a different distribution, which can be described as flat over a redshift range $(0, z_{\text{max}})$. Here I considered $z_{\text{max}} = 5$ for outliers. The probability distribution of the point estimator incorporating systematic errors (SYSPE) can thus formally be written as

$$\text{SYSPE}(z) := \eta \times \text{U}(0, z_{\text{max}}) + (1 - \eta) \times \text{N}(z_{\text{phot}}, \sigma) \quad (4.7.3)$$

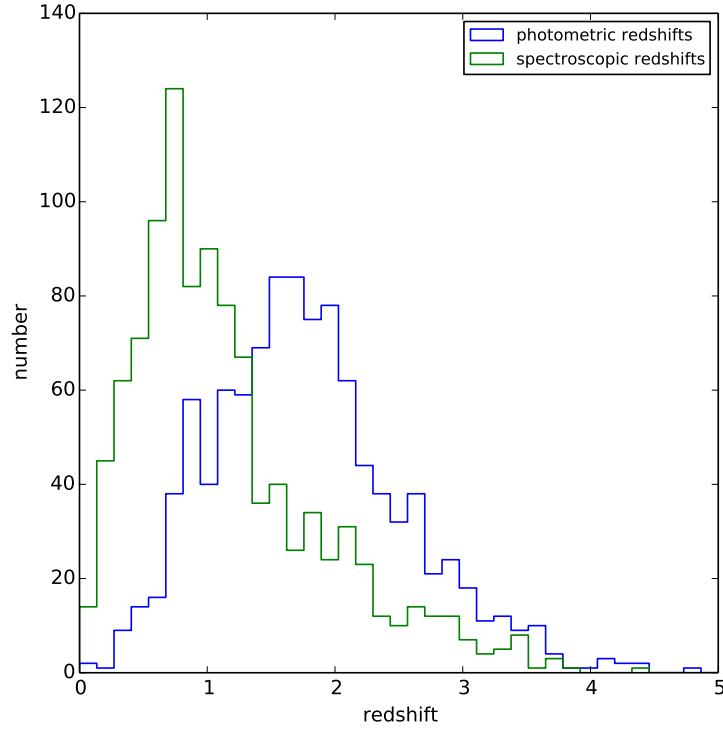


Figure 4.7.1: Redshift distribution of the full sample. Here, the spectroscopic redshifts (green) are separated from the photometric redshifts (blue). The latter reach out also to higher distances.

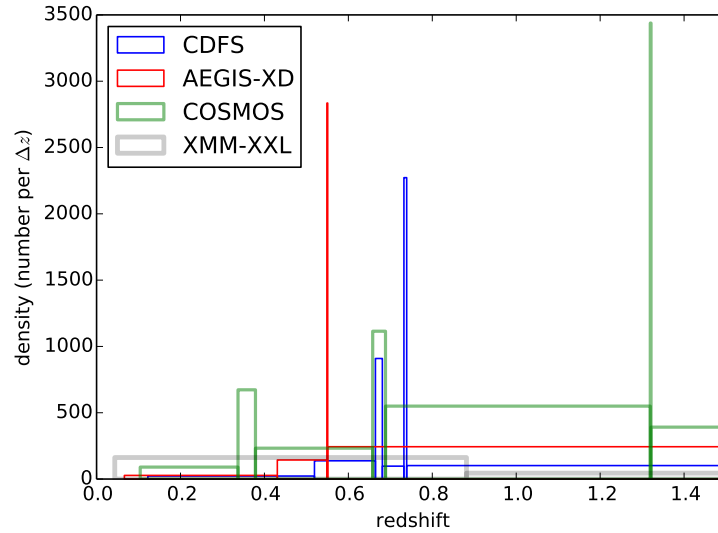


Figure 4.7.2: Redshift distribution in each field. Here, the adaptive histogramming method “Bayesian Blocks” has been used, which splits blocks whenever a significant difference in density is detected. This is both a visualisation and a statistical test for non-uniformity of the redshift distribution. Here, several redshift ranges with over-densities have been identified. Most notable are the peaks at $z \approx 0.5 - 0.7$, which is the age when galaxy clusters are detectable in X-rays. Proximity on the sky is not taken into account in this plot.

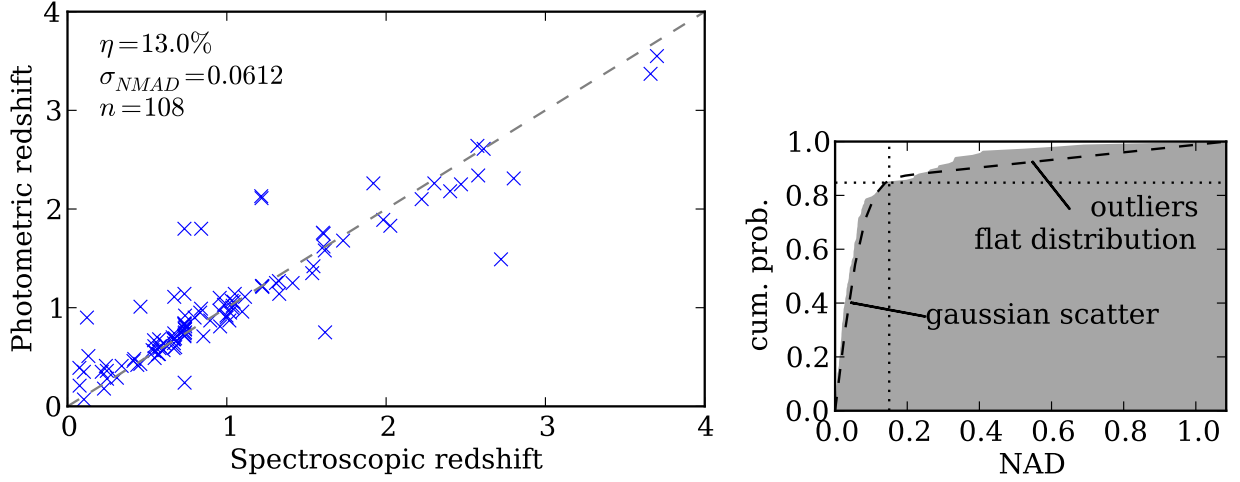


Figure 4.7.3: Photo- z error analysis. The *left panel* shows the distribution of the photometric redshift point estimate for the sub-sample with spectroscopic redshifts analysed in Hsu et al. (2014). The deviation from the 1:1 relationship (grey dashed line) can be described by the key numbers σ_{NMAD} and η . The *right panel* shows that these can model the probability distribution as a combination of a Gaussian and a uniform distribution. Here, the deviation from the 1:1 line, $\text{NAD} = |z_{\text{phot}} - z_{\text{spec}}| / (1 + z_{\text{spec}})$ is considered. The grey area shows the cumulative distribution of the sample. The dashed black line shows the combination of a Gaussian scatter and a uniform distribution based on σ_{NMAD} and η (see text). This approximates the empirical distribution reasonably well.

Here, $U(a, b)$ represents a uniform distribution between a and b , and $N(\mu, \sigma)$ a normal distribution centered at μ with width σ . The right panel of Figure 4.7.3 shows that this modelling in fact matches the empirical distribution very well when considering the cumulative distributions of $|z_{\text{phot}} - z_{\text{spec}}| / (1 + z_{\text{spec}})$.

In this work, I would like to similarly incorporate catastrophic outliers. These are not contained in the photometric redshift distribution (PDZ) from SED fitting. Thus, I follow an analogous modelling of the uncertainty contribution: I allow a broadening of the PDZ by convolution with a Gaussian kernel of width $\bar{\sigma}$ and add a flat probability plateau with weight $\bar{\eta}$:

$$\text{SYSPDZ}(z) := \bar{\eta} \times U(0, z_{\text{max}}) + (1 - \bar{\eta}) \times (\text{PDZ} * N(0, \bar{\sigma})) \quad (4.7.4)$$

The effect of such a convolution on the PDZ is illustrated in Figure 4.7.4. Note that for the special case of $\bar{\sigma} = 0$ and $\bar{\eta} = 0$, SYSPDZ is exactly PDZ. The two parameters $\bar{\sigma}$ and $\bar{\eta}$ have a slightly different definition than σ and η above, but take the same roles for characterising the systematic uncertainty. I fit for the parameters using the spectroscopic sub-sample by maximising the likelihood

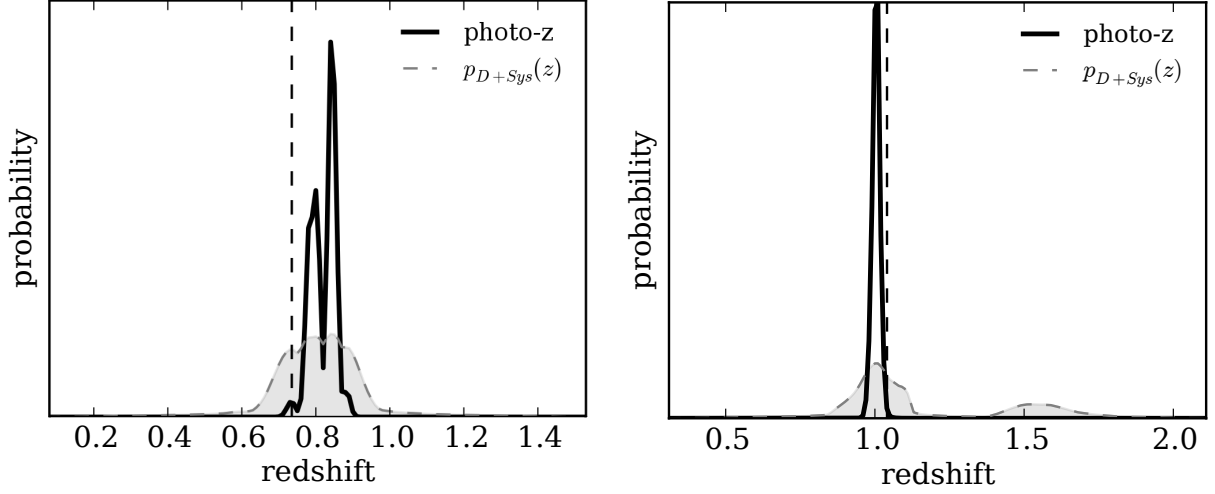


Figure 4.7.4: Demonstration of the kernel convolution of the PDZ. Left panel: Using a smoothing kernel, the initially very sharp redshift distribution is smoothened out. The right panel shows another PDZ that has two solutions, which become more apparent after the smoothing.

$$\mathcal{L}(\bar{\sigma}, \bar{\eta}) = \prod_i \text{SYSPDZ}_i(z_{\text{spec}, i}) \quad (4.7.5)$$

defined as the product of the value of the modified PDZ at the true redshifts z_{spec} . PDZs that are far off the true value will demand an increase in $\bar{\sigma}$ and $\bar{\eta}$, however raising both values diminishes the SYSPDZ value overall due to normalisation of probability distributions.

The best fit parameters are $\bar{\sigma} = 0.024, 0.048, 0.029$ and $\bar{\eta} = 1.8\%, 0.6\%, 2.3\%$ for the CDFS, AEGIS-XD and C-COSMOS samples respectively. These numbers are not directly equivalent to the point estimate based definition of σ_{NMAD} and η (see Salvato et al. (2011) and Hsu et al. 2014), but the numbers are of comparable magnitude. I thus use the smoothed SYSPDZ distributions instead of the PDZs for the analysis in Chapter 6.

A simpler method for incorporating the systematic uncertainty would be to use the modelling based on the point estimator (SYSPE). But when comparing the likelihood value, as in Equation 4.7.5, I found that SYSPE (Equation 4.7.3) always has lower likelihood values than the kernel smoothing SYSPDZ (Equation 4.7.4), even with optimal parameters. This shows that the PDZ contains valuable information due to e.g. secondary solutions, which is missing in the point estimator z_{phot} , and that SYSPDZ is the most faithful representation of the state of the redshift information.

4.7.2 X-ray spectroscopic redshifts (XZ)

Above, I have explained the association of X-ray sources to infrared/optical counterparts in detail, and discussed the computation of photometric redshifts via SED fitting. I have also discussed the systematic errors associated with the fitting method, by comparing a spectroscopic reference sample to the results obtained by photometric redshift estimates. However, one of the remaining sources of uncertainty is the association with optical/IR counterparts, i.e. the problem of confidently choosing the wrong counterpart. The magnitude of this error is also difficult to estimate, especially in the deep fields. One way to improve in this regard is to compute photometric redshift distributions for every possible association, and to marginalise the redshift probability distribution using the posterior probability of the association. This is of course not very practical, due to the unmanageable number of computations involved. Instead, it would be ideal to side-step the association and obtain redshifts from the X-ray spectrum directly. Here, I present the properties of such a technique.

The technique of X-ray redshifts for AGN (XZ) is in principle very similar to SED fitting. A model with a number of parameters, which include the unknown redshift, is compared to the observed data. The details of the technique of applying a model to X-ray data in a Bayesian framework is presented in Chapter 5. It may thus be enlightening for the reader to revisit this section after reading Chapter 5. Without going into detail here, the outcome of the analysis is a probability distribution on the redshift, by marginalising over all other parameters. There are differences to SED fitting however, most notably the wavelength range used. In XZ, the X-ray spectrum (0.5 – 8keV) is used directly. In contrast to photometric redshift fitting, XZ is a spectroscopic redshift technique, although the resolution of the detector is coarse in comparison to optical spectra.

For the current discussion it is sufficient to consider the model to be a photo-electrically absorbed powerlaw. The model is very similar to the `torus+scattering` model presented in Chapter 5, consisting of a powerlaw emitter within a toroidal obscurer (Brightman & Nandra 2012) with an additional scattering component. The scattering component’s power law index and the intrinsic power law are linked, but left free between $1 < \Gamma < 3$. I use uniform priors on all variables except for scale variables like normalizations and the hydrogen column density, for which I use a Jeffreys prior (same as in Chapter 5, see motivation there). For the redshift, a flat prior between $0.1 < z < 6$ is used. Finally, I added a collisionally ionised plasma component (`apex`) to model stellar contributions from the host galaxy.

First I demonstrate that the method is clearly capable of constraining the redshift, and discuss why this is possible at all. Figure 4.7.5 and 4.7.6 shows the posterior on the redshift parameter obtained for four objects in the CDFS, using the X-ray spectrum only. The corresponding X-ray spectra are shown in Figure 4.7.5. The top spectrum, ID 430, shows a virtually flat spectrum, indicating high column density. The Fe-K α line is visible, and presumably this is the feature constraining the redshift in this spectrum. The second spectrum (ID 87) shows a very different case. Here, the overall shape matches the data, and both exhibit several “wiggles”. These are due to the detector’s non-linear

behaviour, which is of course independent of the redshift of the source, and the absorption edges of the photo-electric absorption, which are dependent on the redshift parameter. By varying the parameters, the XZ method found that a single, well-constrained combination of redshift, obscuring column density and normalisation matches the shape of the spectrum better than any other. This allows the redshift to be constrained. Figure 4.7.6 shows two more examples. The spectra indicate that several features in the X-ray spectrum can help constrain the redshift, most notably the edges in the photo-electric absorption cross-section and the Fe-K α line.

Now I have established that the XZ method obtains the correct redshift for at least some sources. But spectra of AGN exhibit a wide diversity of features in the 0.5 – 8keV band, as shown in Figure 4.7.5. For the technique to succeed, it is important for the model to capture all of these effects, as a mismatched model specification can result in catastrophic redshift estimates, and unwittingly introduce large systematic uncertainties in the derived spectral parameters. Thus, to understand if this is a problem, the next step is to investigate under which circumstances the XZ method delivers good results, and how often it gives incorrect results.

Figure 4.7.7 shows that the obscuring column density and the number of detected counts have influence on the amount of information gained for the redshift parameter. The other spectral parameters such as the photon index Γ , scattering component, the redshift, etc. do not show any trends with the information gain. In other words, only the number of counts and the obscuration determine whether the redshift can be constrained. Here, the information gain (IG) is used as a measure how much the posterior distribution differs from the prior (see Section 5.2.4 for a formal definition). In the case here it just measures how concentrated the PDZ is (see Figure 4.7.5 and 4.7.6 for illustration).

To test whether the derived redshift is correct, a sample of objects with secure spectroscopic redshifts is selected from the CDFS. 295 sources were considered. 78 sources (26%) show significant information gain (> 1.2 nats) using the XZ technique. These are shown in Figure 4.7.8. For all but a few sources (6% in the CDFS), the true redshift is bracketed by the XZ estimate. This fraction of outliers is comparable to the outlier fraction of photometric redshift techniques (see above).

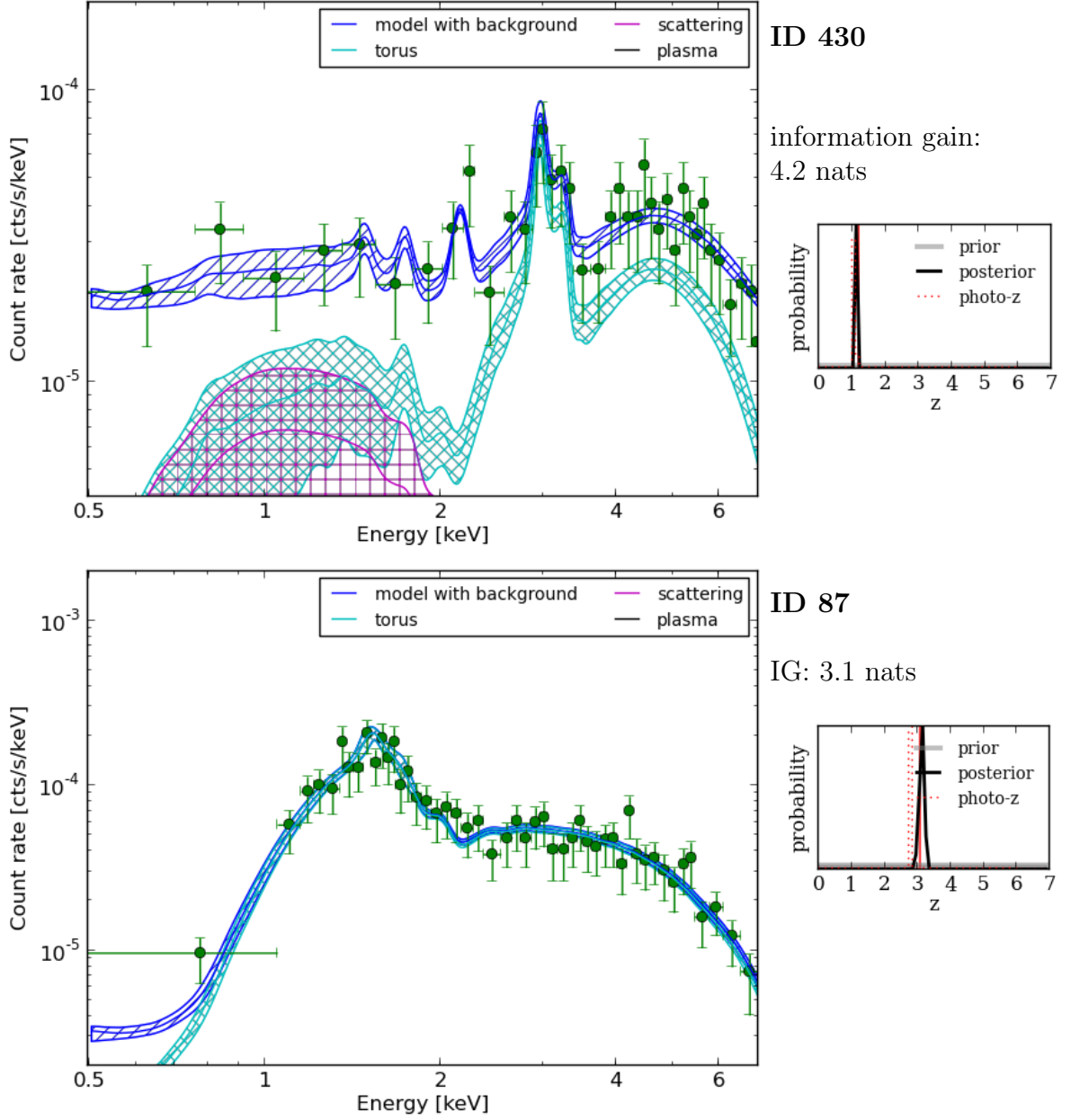


Figure 4.7.5: Redshift constraints obtained using the XZ method. For four selected objects (continued in Figure 4.7.6), the X-ray spectrum (left panels) and the obtained redshift posterior distribution is shown (right panels). In the *right panels*, the source ID and the information gain is given. The information gain is a measure of the difference between the flat redshift prior distribution and the posterior. Sources with good constraints show $KU > 1.2$ nats. For comparison, the photo-z estimate (PDZ, dotted red) and the spectroscopic redshift (vertical solid red line) are shown when available. The corresponding spectra are shown in the *left panels*. In green, the obtained counts are shown, binned for visualisation. Blue shows the spectral model. The range indicates the 10% and 90% quantiles. (see text)

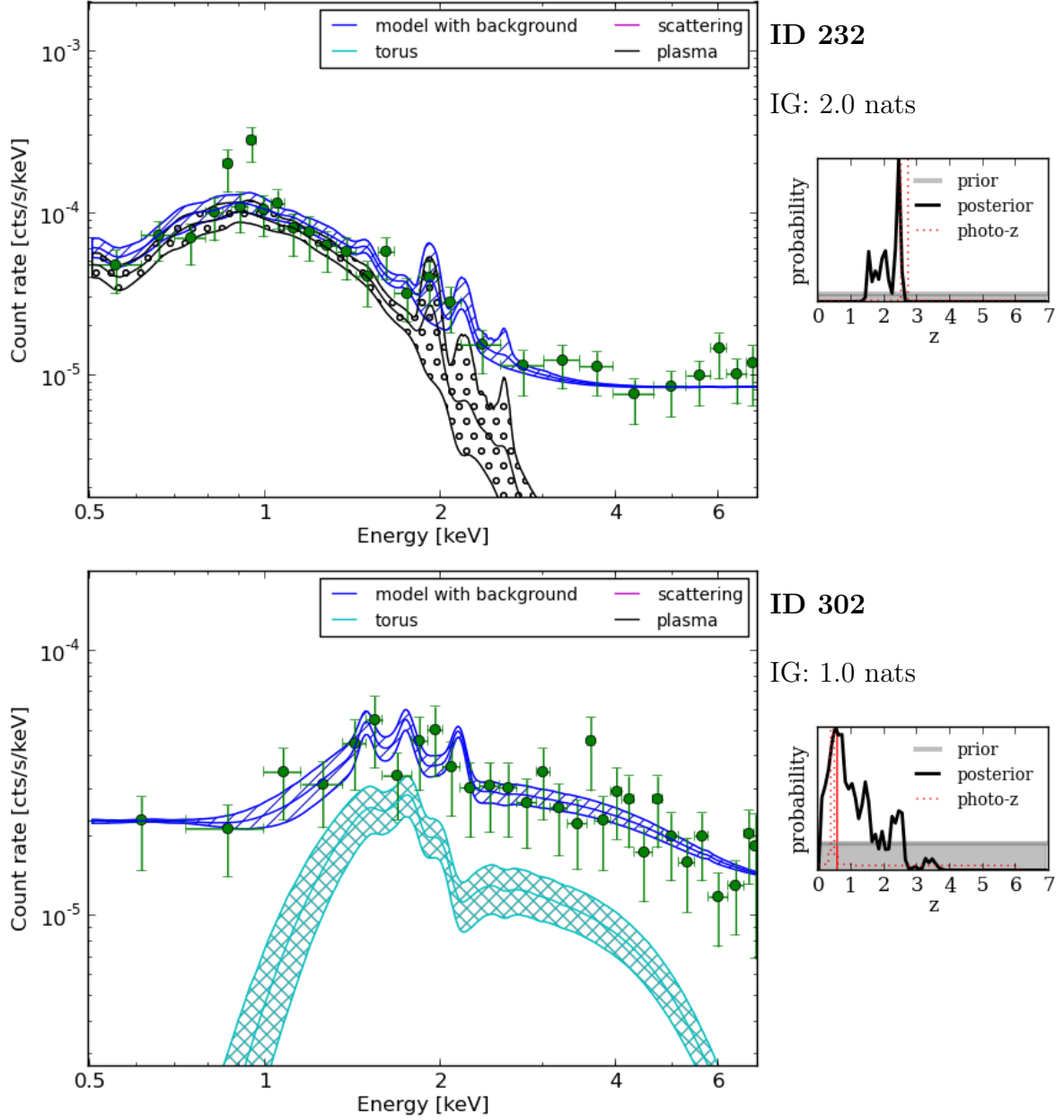


Figure 4.7.6: Continuation of Figure 4.7.5. The top spectrum (ID 232) shows a case where the spectrum of the AGN is mixed with stellar contributions of the host galaxy. Here it is probably the stellar component that constrains the redshift, although the redshift posterior is very broad (low information gain, IG). In the bottom spectrum (ID 302) I show a spectrum that shows few features. Essentially, only redshifts at $z > 2.5$ could be excluded by the XZ technique.

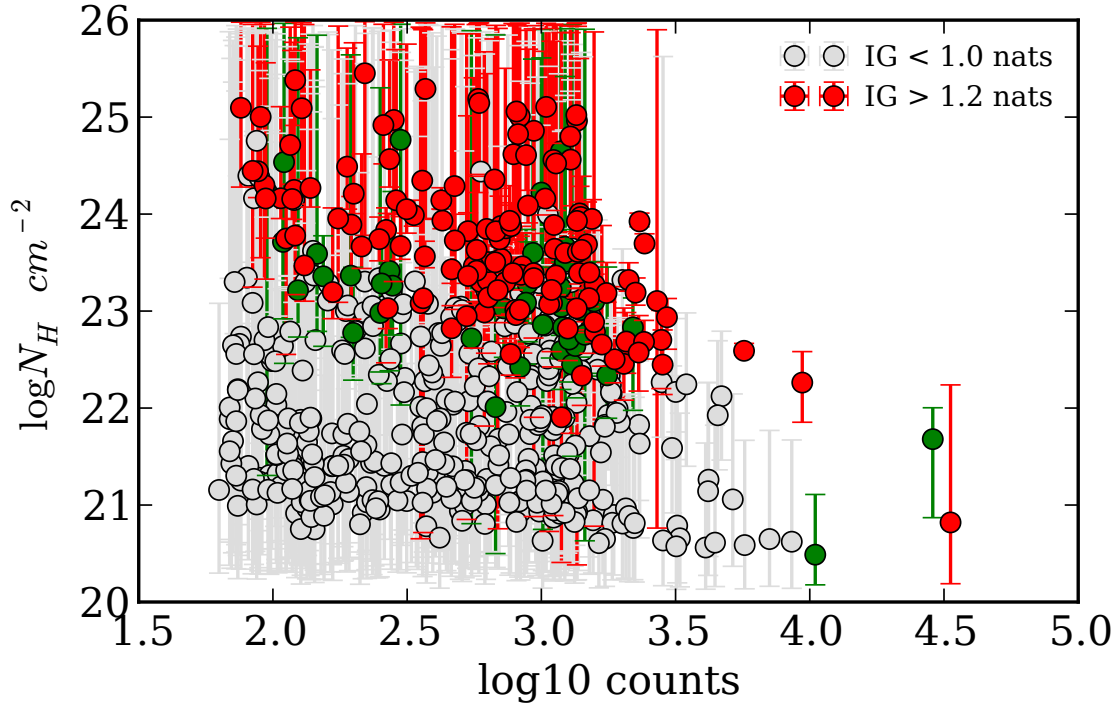


Figure 4.7.7: Dependence of information gain (IG) on column density and number of X-ray counts. Sources with $\text{IG} > 1.2 \text{ nats}$ (red points) have good redshift constraints (see 4.7.5 for illustrations). The remainder are plotted in gray (less than 1.0 nats) and green (intermediate). The required number of counts and column density to obtain good constraints are shown. For sources with $N_H > 10^{23} \text{ cm}^{-2}$, 100 counts are required for good constraints. For sources with lower column density ($N_H = 10^{22-23} \text{ cm}^{-2}$), at least 1000 counts are necessary. For unobscured sources ($N_H < 10^{22} \text{ cm}^{-2}$), it is generally not possible to constrain the redshift. This shows that the determining factor for good redshift constraints is the AGN obscuration, via absorption edges.

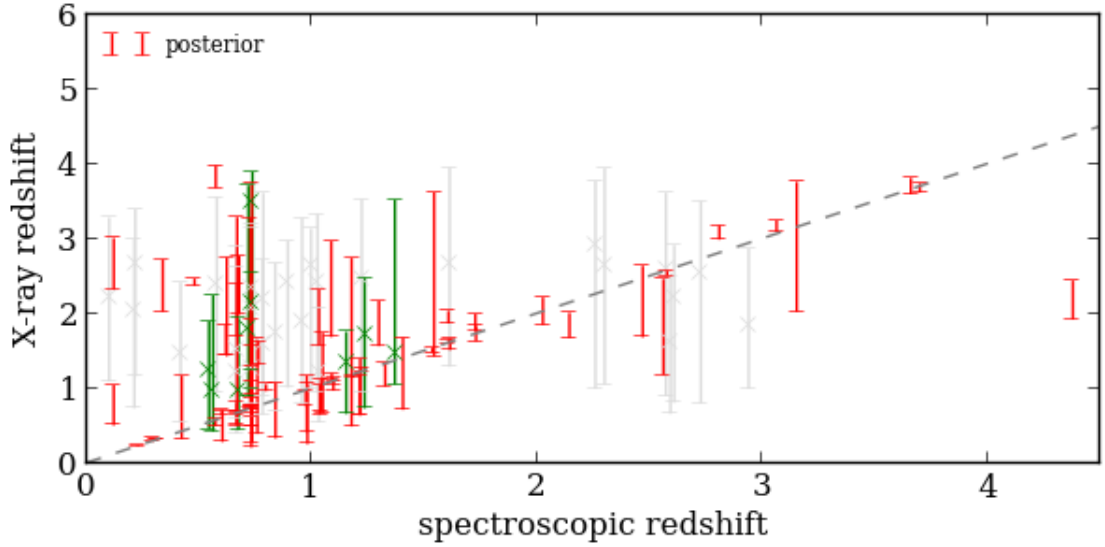


Figure 4.7.8: Evaluation of XZ redshift estimates. The dashed, grey line shows the ideal 1:1 relationship. The error bars indicate the 10% and 90% quantile range of the XZ estimate. Sources with significant information gain, $IG > 1.2$ nats are shown in red (see 4.7.5 for illustrations). The remaining sources are plotted in gray (less than 1.0 nats) and green (intermediate).

In conclusion, the XZ method of obtaining redshifts directly from the X-ray spectrum can constrain the redshift tightly, with a low fraction of outliers. This method effectively side-steps the costly and difficult association and redshift estimation using infrared and optical wavelengths. However, this method is only applicable in a minority of sources ($\sim 25\%$ in the CDFS) with sufficient constraints. There, however, the error rate is low, indicating a powerful complementary approach that can assist optical/infrared redshift estimation techniques.

4.8 Sample selection

I now return from these side-notes on redshift estimation to the treatment of the X-ray spectra, which prompted redshift estimation in the first place. As a last preparation step, I remove spectra that may be contaminated by non-AGN processes.

Based on the initial hard-band detection, passive (non-AGN) galaxies should already be absent from the sample, except for nearby, star-forming galaxies. The latter can be removed to a large degree by excluding intrinsically faint sources that have $L_X < 10^{42}$ erg/s with at least 90% probability. Additionally, stars have been removed during redshift estimation (see Section 4.7). The number of objects used in the luminosity function analysis (Chapter 6) is shown in Table 4.2. Figure 4.8.1 visualises the properties of the sample.

For the study of the obscurer in Chapter 5, more stringent criteria are necessary. There, I am studying the involved processes in detail, instead of only estimating parameters such

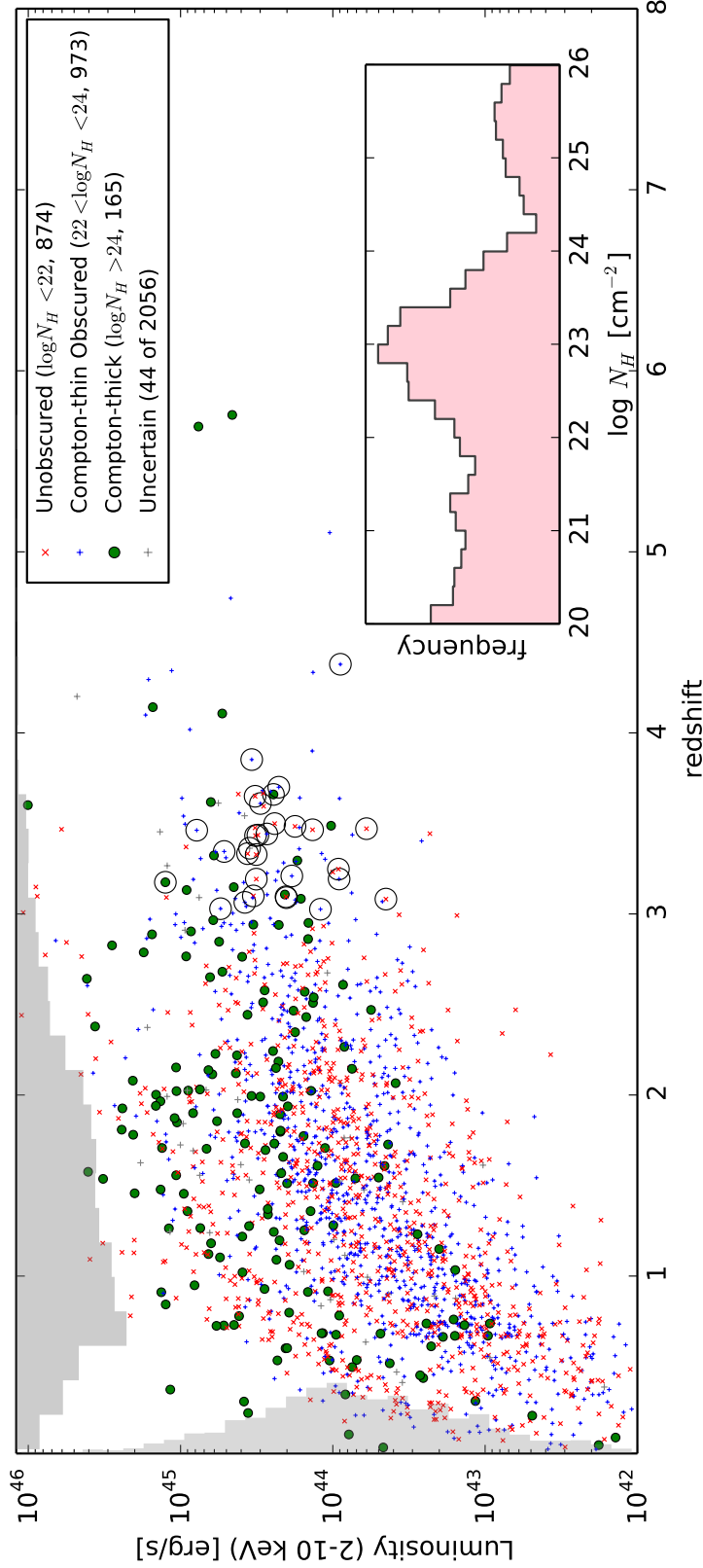


Figure 4.8.1: Luminosity-redshift plot of the full sample. The intrinsic 2 – 10 keV luminosity in erg/s and redshift is plotted. For this visualisation, sources are coloured based on $\log N_H$ (in units of cm^{-2} as follows: 20 – 22, “Unobscured”, red; 22 – 24, “Compton-thin Obscured”, blue; 24 – 26, “Compton-thick”, green). If more than 50% of the N_H posterior probability lies within one of the intervals above the source is colour-coded accordingly. Due to the heavy suppression of the flux of Compton thick objects by absorption, they are typically only detectable at higher intrinsic luminosity compared to unobscured or Compton-thin sources. Some objects lie below the $L = 10^{42} \text{erg/s}$ limit and are not used in this work. The inset shows the N_H histogram of the sample, while the top and left axes show histograms of the redshift and luminosity distributions in gray. These plots are constructed by drawing a random posterior sample for each object, and thus includes the uncertainty in the parameter estimation as well as the intrinsic distribution. Sources above redshift $z = 3$ which have spectroscopic redshift estimates are indicated with large black circles.

as the column density. These exclusion criteria are described in the next section.

4.9 Sub-sample for study of the obscurer geometry

For the study presented in Chapter 5, it is not important to have an unbiased sample of AGN. Instead, a clean sample is required that only includes AGN contributions, ideally in all their variety. This study was performed based on the CDFS survey.

Star-burst galaxies can contaminate the sample and skew the inferences about AGN, especially with regards to the soft energy ranges. I adopt the criteria for inclusion of AGN from Xue et al. (2011): Sources with $L_{X, 2-10\text{keV}} < 3 \times 10^{42}\text{erg/s}$, effective power law index $\Gamma_{\text{eff}} > 1$ and $\log L_X/L_{\text{opt}} < -1$ are not used. These criteria select weak X-ray sources that are much brighter in the optical, and additionally emit mostly soft X-rays. This selection is designed to exclude sources with non-negligible host contribution. Thus, moderate-luminosity AGNs in star-burst galaxies and low-luminosity AGN are not studied. Furthermore, objects classified as stars or galaxies (i.e. host-dominated sources in the X-ray) by Xue et al. (2011) are removed. From the initial sample of 526 X-ray sources detected in the CDFS in either the soft or hard band, 346 AGN remain.

Model selection assumes that one model is the correct one (see Section 5.2.3). For determining whether the model could produce the data at hand, I adaptively bin the spectrum counts so each bin contains 10 counts. Then I compute the χ^2 -Statistic for the best fit model parameters. If $\chi^2/n > 2$, where n is the number of bins, the object is not used for model comparison. For low count spectra, this criterion is relaxed further (2.3 if less than 500 counts, 3 if less than 100 counts, 5 if less than 50 counts) due to the stronger Poisson variance. These limits were obtained by simulating a flat Poisson spectrum across bins, so that they exclude the true value in fewer than 1% of the simulations. The 13 affected sources were visually inspected in the X-ray and optical and at least 5 of them can be clearly explained by outliers in the photometric redshift due to contaminated photometry or incorrect association. With an outlier fraction of $\eta = 4 - 6\%$, the expected number of outliers is ~ 20 . 334 AGN remain in the sub-sample after these objects are excluded.

5 The nature of the obscurer around SMBH

A striking result from observational data in the last 30 years show that a large fraction, if not the majority of AGN, are obscured by cold, molecular material that blocks a direct view of central AGN. One central question in understanding the entire AGN population is thus to study the fraction of obscured AGN and the basic geometrical properties of the associated gas and dust clouds. In order to robustly characterise the obscuring material, it is essential to understand the effects of this line-of-sight obscuration. A number of studies have investigated the column density distribution using different methodologies, including hardness ratios and spectral models of varying complexity.

Studies of nearby, obscured AGN have demonstrated complex spectra that include a photo-electrically absorbed powerlaw, a soft X-ray excess which can be associated to an ionisation cone, strong $\text{FeK}\alpha$ line emission, ionised bulk outflows, and/or a strong reflection contribution (see Introduction section 3.2). These effects indicate that the gas is, at least in these objects, arranged in a complex way. In order to get robust constraints on the obscuration all these physical processes need to be accounted for.

In deep field observations, which preferentially detect high-redshift objects, simpler techniques have been applied due to the limited count statistics. For example, Cappelluti et al. (2009) and La Franca et al. (2005) measure hardness ratios (colors between X-ray bands) of XMM observations to determine whether deep field AGN are obscured. In the case of a photo-electrically absorbed powerlaw of known slope, the hardness ratio provides a measure of the obscuring column density. At a certain point however, the hardness ratio saturates. The hardness ratio is problematic in more complex spectra, as e.g. the presence of the soft component, Compton scattering and the $\text{Fe K}\alpha$ feature do not allow a unique determination of the obscuring column. Furthermore, the background can not trivially be subtracted in the Poisson regime of low counts.

Other studies have employed spectral fitting with more advanced spectral models to capture these effects. For instance, Tozzi et al. (2006) uses an absorbed powerlaw with a soft powerlaw component of the same spectral index and a Compton reflection spectrum. Evidence is found for the soft and/or reflection component in 8 and 14 CDFS sources respectively (of a total of 347). In this and other studies, the detection of spectral components is based on likelihood ratios using a fixed threshold (essentially motivated by Wilks' theorem, see Section 5.2.2). The Compton scattering and absorption are modelled independently, without considering FeK line fluorescence.

Studies based on hardness ratios (e.g. La Franca et al. 2005) and those using X-ray spectral fitting (Ueda et al. 2003; Hasinger et al. 2005; La Franca et al. 2005; Treister &

Urry 2005; Tozzi et al. 2006; Akylas et al. 2006; Ueda et al. 2014) have investigated the N_H distribution and its dependence on redshift and/or X-ray luminosity. It is worth noting that sources are typically treated differently depending on the number of counts to avoid systematics introduced by overfitting, such as multiple solutions in a complex spectrum. Such systematics could propagate into subsequent analyses (e.g. luminosity functions, N_H distribution; discussed in Chapter 6).

Brightman & Ueda (2012) improved the spectral modelling by considering physically motivated, toroidal shapes for the obscurer. Their Monte Carlo simulations (Brightman & Nandra 2011a) consistently model e.g. photo-electric absorption, Compton reflection and $\text{FeK}\alpha$ emission. Depending on the source they find evidence for different torus models, which include a torus with opening angles of 60° , 30° and 0° (sphere), each in comparable fractions of the sample. In this study a fixed value for the redshift is used, which in the majority of sources comes from photometric redshift methods. For comparing models which introduce a new parameter, Brightman & Ueda (2012) use a likelihood threshold of $\Delta C > 1.0$, corresponding to the fact that the new parameter excludes the simpler model within its 1σ region. For models with equal number of parameters, e.g. when comparing torus models with different opening angles, the model with the highest likelihood is accepted.

This chapter presents substantial methodological advances for spectral analysis of AGN and the modelling of physical effects in deep X-ray observations. To this end, the spectra of sources detected in the CDFS (see Section 4.9) are analysed. A total number of 10 physically motivated models for the structure of AGN are tested (presented in Section 5.1). The most notable contribution is the use of Bayesian model comparison to compare whether increasingly complex spectral models are justified. Furthermore, the uncertainty associated with photometric redshift estimates (see Section 4.7) can incur degeneracies with N_H estimates (see Section 5.3.1). In the Bayesian methodology presented here, the photometric redshift uncertainty is propagated into the derivation of spectral properties. These effects have not been addressed in previous studies but are potentially important. The following chapters which investigate the column density and luminosity distribution of the AGN population (Chapter 6 and 7) benefit from the determined spectral parameters and their realistic uncertainties. The use of a Bayesian methodology circumvents the problem of overfitting and allows us to use physically realistic, complex spectral models also in the low count regime where multiple solutions occur. Thus the new methodology allows robust decisions under the limited data collected and can be applied consistently both in the low and high-count regime. The statistical methodologies are discussed in detail in Section 5.2.

This work was published as Buchner et al. (2014) in *Astronomy & Astrophysics*, Volume 564, id. A125, 25 pp.

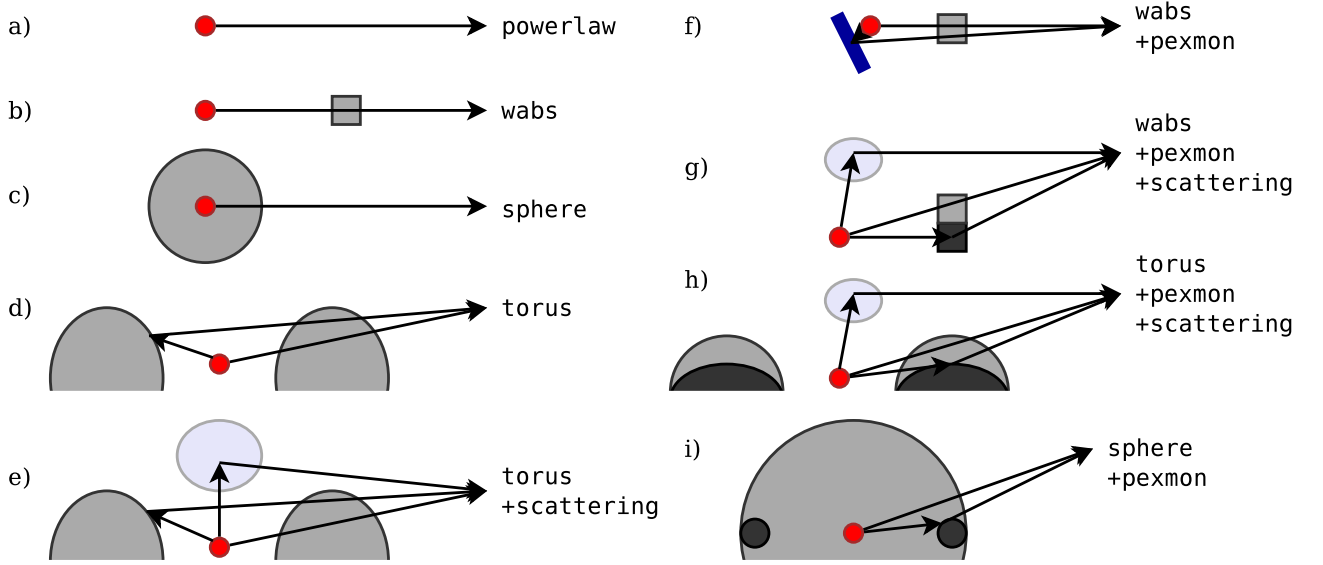


Figure 5.1.1: Cartoon illustrations of the geometries associated with each model. The **wabs** model (b) represents an absorbing slab in the line of sight, and can also be interpreted as the case of a torus with extreme opening angle. While **torus** (d) uses an intermediate opening angle, the **sphere** (c) represents the other extreme, a vanishing opening angle. The **+scattering** extension (e) of the named models correspond to Thomson scattering from outside the line of sight, which is not subject to any absorption. Finally, the reflection component (**+pexmon**) corresponds to either disk reflection (f) or additional reflection if the torus is not viewed through the same column density as the reflection (g, h, i). For the **sphere** it should be noted that scattering is not physically possible, as no unabsorbed radiation can escape.

5.1 Model definitions

The focus of this chapter is a study of the AGN population to constrain the geometry of the X-ray obscuring material in the vicinity of SMBHs.

Above, I have discussed the mechanisms emitting the “intrinsic” powerlaw spectrum, due to the accretion disk and corona. This spectrum can be modified substantially by surrounding gas, absorbing and reprocessing the primary radiation. To identify the physical mechanisms of this obscurer and its geometry I describe and compare 10 physically motivated models with different levels of complexity. For this model comparison, I developed a Bayesian methodology (discussed in Section 5.2) that propagates the uncertainties from X-ray observations and redshifts correctly. The data used is the obscurer sub-sample described in Section 4.9, selected to be free of host galaxy contribution (e.g. by stellar emissions). The models and their spectral effects considered are defined below.

The simplest model considered is a power law, referred to as **powerlaw**. It represents the intrinsic spectrum of AGN without any obscurer, as shown in Figure 5.1.1a). The high-energy cutoff observed in the energy range 80 – 300 keV (Perola et al. 2002) lies well above the energy range used in this work (< 10 keV) even at moderate redshifts. I therefore

Symbol	Description	Unit	Prior	Introduced in
	Normalisation	cts/s	Log-Uniform between 10^{-30} and 1	powerlaw
Γ	Photon index	-	Normal distribution $N(\mu = 1.9, \sigma = 0.15)$	powerlaw
N_H	Column density	cm^{-2}	Log-Uniform between 10^{20} and 10^{26}	wabs, sphere, torus
f_{scat}	Soft scattering fraction	-	Log-Uniform between 10^{-10} and 10^{-1}	wabs+scattering, sphere+scattering, torus+scattering
R	Reflection fraction	-	Log-Uniform between 10^{-1} and 10^2	wabs+pexmon+scattering, sphere+pexmon+scattering, torus+pexmon+scattering

Table 5.1: Parameters of the models used.

neglect the cutoff. The parameters of **powerlaw** are the normalisation at 1 keV and the power law index, Γ (see Table 5.1).

The most commonly used model to describe obscuration is to apply photo-electric absorption to the intrinsic power law. Absorption has the largest cross-section among the interaction processes, and thus is a good first order approximation for the X-ray spectra of AGN detected in deep fields. However, at higher column densities ($N_H \gtrsim 10^{24} cm^{-2}$), matter becomes optically thick to Compton scattering, and re-emission in lines due to X-ray fluorescence is prevalent. In contrast to photo-electric absorption, which is opaque at low energies, Compton scattering introduces an energy loss at each interaction, thus low-energy photons can be received by the observer. Furthermore, Compton scattering induces a non-uniform scattering angle distribution making the spectrum dependent on the assumed geometry. Influenced by the solid angle illuminated by the source and the surface area exposed to the observer, the contribution by Compton scattering varies, as radiation is scattered into the line of sight. For instance, a torus geometry with a certain opening angle produces a smaller Compton reflection hump between $\sim 10 - 30$ keV (see Figure 3.2.1) than a completely closed, spherical geometry (Murphy & Yaqoob 2009; Brightman & Nandra 2011a). The opening angle of the torus, and the viewing angle to a minor degree, thus influences the strength of the reflection hump. This allows one in principle to determine the viewing/opening angle and N_H independently, possibly probing the density gradient of the obscurer. However, because the effects on the observed spectrum are small, particularly in the case of low photon count spectra, this technique has not yet been successfully applied to discriminate these parameters in individual obscured sources. In this work, a limited range of geometries is assumed and I do not attempt to constrain the viewing angle. The geometry-dependence of spectra also makes modelling challenging. Multiple inter-dependent interactions of several elements can realistically only be done by X-ray radiative transfer simulations for a fixed geometry (see e.g. Nandra & George 1994;

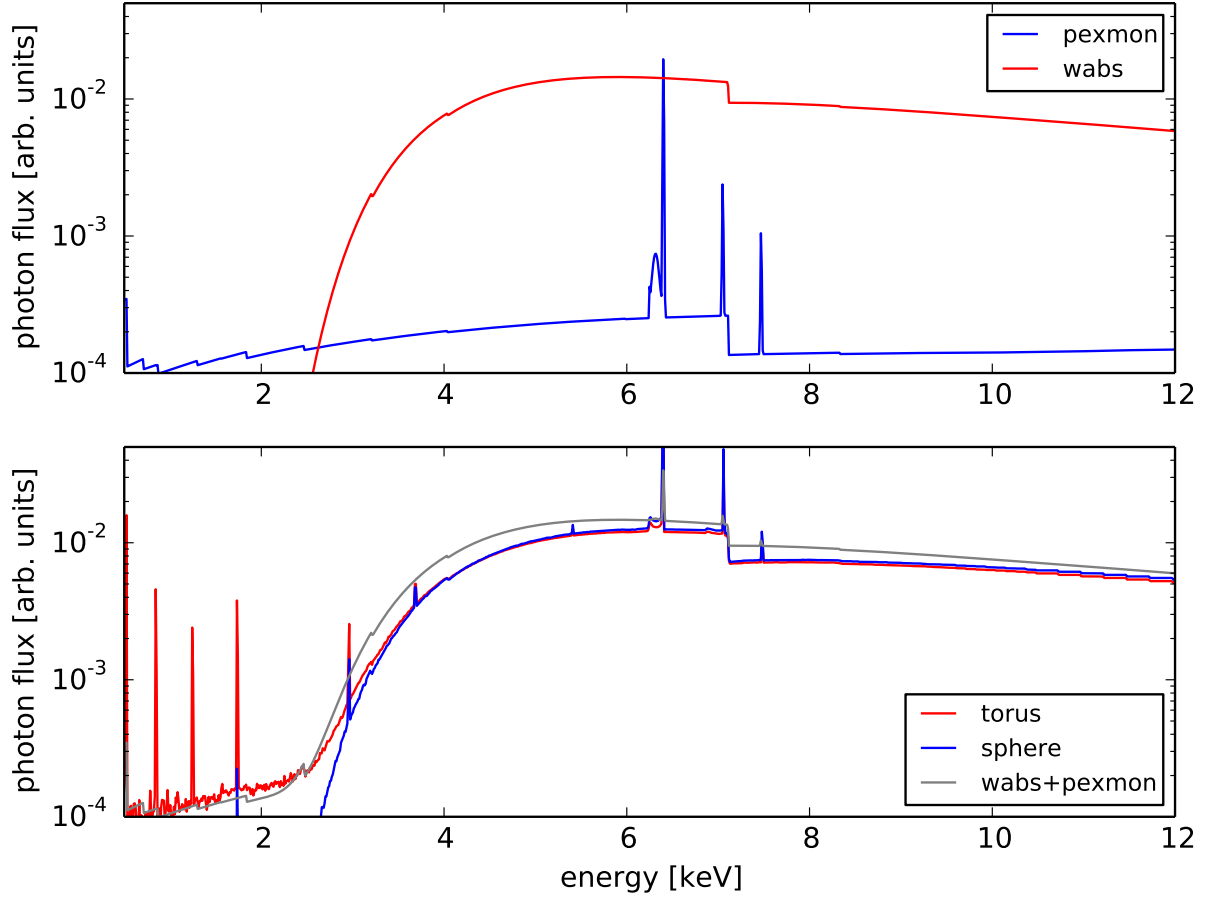


Figure 5.1.2: Illustration of model spectra. Here, an intrinsic powerlaw with $\Gamma = 2$ is absorbed by various obscurer models, always with a column density of $N_H = 10^{23.5} \text{ cm}^{-2}$. *Top panel*: The **wabs** model (red line) shows the effect of photo-electric absorption only. The **pexmon** component (blue line) illustrates the contribution of Compton scattering and fluorescence, as it models the reflection off a optically thick slab (with $R = 0.1$, see text). In the *bottom panel*, their combination is shown as **wabs+pexmon**. This combination is nearly indistinguishable from self-consistent modelling of Compton scattering, photo-electric absorption and iron K-shell fluorescence effects, as done in the **torus** geometry (red line). The **torus** model provides more Compton reflection than the **sphere** geometry (blue line), because the backside of the torus can act as a mirror. On the other hand, the **sphere** model shows slightly more fluorescence and a slightly Compton hump at high energies (note that photons are absent at low energies).

Murphy & Yaqoob 2009; Brightman & Nandra 2011a).

In this work, three obscuration models are considered, **wabs**, **sphere** and **torus**, which correspond to different geometries:

The “**wabs**” model applies only photo-electric absorption to a power law with the cross-sections and polynomial approximations computed by Morrison & McCammon (1983). This model does not include any Compton scattering. One physical scenario consistent with this model is an infinitely small blob in the line of sight (LOS), as shown in Figure 5.1.1b). In this scenario, no radiation is scattered into the LOS, and all Compton scattering leaves the LOS (although this loss is not part of the model), and thus the spectrum only consists of photo-electric absorption.

I also consider an absorber with spherical geometry (**sphere**) illustrated in Figure 5.1.1c. This model, computed by Brightman & Nandra (2011a), self-consistently models photo-electric absorption, Compton scattering and K-shell fluorescence of a power-law spectrum source at the centre of a cold, neutral medium. Similarly, a toroidal geometry (**torus**) is used, which was simulated in the same way as **sphere**, but with a bi-conical cut-out. I define the torus model to have a 45° opening angle viewed edge on. The physical scenario represented by this model is shown in Figure 5.1.1d, which indicates that Compton scattering into the LOS is possible.

The three obscuration models used mark the extreme cases of torus geometries: **wabs** for an almost 90° half-opening angle, **sphere** for a vanishing opening angle and **torus** represents an intermediate case where 30% of the incident radiation encounters the obscurer. Figure 5.1.1b-d illustrates these differences. In comparison to the **powerlaw** model, **wabs**, **sphere** and **torus** have the additional parameter N_H , the neutral hydrogen equivalent column density in the LOS (see Table 5.1).

Observations of some obscured AGN show that a fraction of the intrinsic radiation escapes without encountering any obscuration. This is understood to be Thomson scattering off ionised material inside the opening of the obscurer, as illustrated in Figure 5.1.1e for the torus geometry (Turner et al. 1997b). I model this component¹, referred to as (soft) scattering (**+scattering**) by adding a simple, unabsorbed power law component with the same Γ as the incident radiation. The normalisation of this component, f_{scat} , is modelled relative to the intrinsic power-law component, and may vary between 10^{-10} and 10%. In this fashion, **torus** and **wabs** are extended to **torus+scattering** and **wabs+scattering**. For the sphere model, it should be noted that no unabsorbed radiation can escape, and thus adding scattering is unphysical. However, for completeness, I also consider **sphere+scattering**, which mimics the case of a torus with a very large covering fraction.

Any obscurer geometry immediately gives rise to a distribution of observed column densities N_H , as the populations objects are viewed from random viewing angles. Treister et al. (2004) argued that the observed N_H distribution requires a stronger gradient than a constant-density torus geometry can provide. Thus an additional density gradient contri-

¹This component can also model other contributions to the soft energy range, such as X-ray binaries and hot gas from galaxies, although we try to remove such sources in the sub-sample used here (see Section 4.9).

bution is considered by adding a denser slab outside the LOS inside the obscurer. The N_H – measured largely through photo-electric absorption – indicates only the “effective” column density along the LOS, and thus remains unaffected. However, a denser region outside the LOS can scatter Compton reflection into the LOS, as illustrated in Figure 5.1.1g-i. Alternatively, reflection off the accretion disk may also be part of the LOS spectrum entering the obscurer (Fabian 1989; George & Fabian 1991). This is illustrated in Figure 5.1.1f for **wabs**. Thus, a Compton reflection component **+pexmon** is added to the obscured spectrum. Here it is assumed that this component passes through the obscuring column, so it is absorbed with the same column density as the LOS obscuring material. The **PEXMON** model (Nandra et al. 2007b) combines Compton reflection off a neutral dense semi-infinite slab geometry (**PEXRAV**, Magdziarz & Zdziarski 1995) self-consistently with $Fe K\alpha$, $Fe K\beta$ and $Ni K\alpha$ emission lines and the $Fe K\alpha$ Compton shoulder (George & Fabian 1991; Matt 2002). The same incident spectrum (photon index Γ and no cutoff) is assumed for the reflection, and I follow Nandra et al. (2007b) in assuming a fixed inclination of 60° . The normalisation of this component is modelled relative to the direct powerlaw emission (R parameter), and may vary between 0.1 and 100. Figure 5.1.1 illustrates various causes for an additional reflection spectrum (f-i). In the case of **wabs** model in particular, the **+pexmon** component can compensate for the lack of forward-scattering inside the obscurer.

To summarise, 10 models are considered, identified by **typewriter** font style. The statistical method used for model comparison is described in Section 5.2. Figure 5.1.3 illustrates the comparisons of interest with arrows. Firstly, obscurer geometries are compared (**powerlaw**: no obscurer, Figure 5.1.1a; **wabs**: bullet-like blob, 5.1.1b; **sphere**: complete covering of the source, 5.1.1c; **torus**: intermediate case, 5.1.1d). Then, the existence of a soft scattering component is tested for, represented by **wabs+scattering**, **sphere+scattering** and **torus+scattering** (Figure 5.1.1e). Finally, the need for additional Compton reflection as shown in Figure 5.1.1 f,g is explored via **wabs+pexmon+scattering**, 5.1.1i for **sphere+pexmon+scattering** and 5.1.1h for **torus+pexmon+scattering**. The following section describes the methodology used to distinguish between the models, and to constrain the model parameters.

5.2 Statistical analysis methods

5.2.1 Parameter estimation

In X-ray spectral analysis it is common practice to vary the parameters of the spectral components until a certain statistic is optimised with regard to the observed X-ray spectrum.

Because X-ray CCD detectors are photon-counting instruments, their spectral bins are filled by counts according to a Poisson process. In faint sources, which constitute the majority of the detections, the number of counts is so low that the Gaussian approximation does not hold. In this regime, the familiar Gaussian χ^2 is replaced by the maximum likelihood C-Statistic $C = -2 \times \ln \mathcal{L}_{\text{Poisson}} + \text{const}$ (Cash 1979), based on the Poisson

likelihood $\mathcal{L}_{\text{Poisson}}$. Both the C -statistic and χ^2 denote the $-2 \times \ln$ of a likelihood.

A Bayesian approach is used for estimating model parameters and their uncertainties. van Dyk et al. (2001) introduces Bayesian X-ray spectral analysis in detail. In a Bayesian approach, the Poisson likelihood can be weighted by priors into a posterior probability distribution. The goal in Bayesian parameter estimation is then to identify sub-volumes which constitute the bulk of the probability integral over the parameter space. This approach performs optimisation and error estimation simultaneously, but requires an integration and exploration technique that performs well in multiple dimensions (i.e. avoids the “curse of dimensionality”).

It is common practise to employ Markov Chain Monte Carlo (MCMC) for this type of problem. However, I use the Nested Sampling technique (Skilling 2004), which has several advantages over MCMC: (1) the integral over the parameter space is computed, which is important for model comparison (see below), (2) The MCMC algorithm requires analysis for convergence, which is not trivial and (3) the parameter space may have multiple solutions in the form of islands in the parameter space, or highly degenerate parameters (e.g. N_H and L). MCMC has to be specially tuned to deal with such scenarios.

Nested Sampling keeps a set of fixed size of parameter vectors (“live points”) sorted by their likelihood. These points are always randomly drawn from the prior distribution. The algorithm then removes the least likely point by drawing points until one is found with a higher likelihood. Effectively, this approach “scans” the parameter space vertically from the least probable zones to the most probable. For each removed point, the volume for which this point is representative of is computed, and the according probability mass (likelihood times volume) added to the integration. When the remaining parameter volume is negligible, the algorithm terminates. A difficulty of this algorithm is to avoid the curse of dimensionality when drawing from the prior distribution to get higher valued points.

MultiNest (Feroz et al. 2009) solves this problem efficiently by clustering the live points into multi-dimensional ellipses, and drawing from these subspaces under the assumption that higher-valued points can only be found in proximity of already drawn points. Because of the clustering, MultiNest can follow distinct local maxima without difficulty. MultiNest is applicable to the low-dimensional problems of X-ray spectral modeling. It can compute points of equal weighting akin to a Markov Chain, provide values and error estimates for each local maximum as well as marginal probability distributions for each parameter. The integral over the parameter space is computed globally and for each local maximum.

5.2.2 Model comparison

Bayesian model comparison is performed by comparing the integrals over parameter space, called the evidence $Z = \int \pi(\vec{\theta}) \exp \left[-\frac{1}{2} C(\vec{\theta}) \right] d\vec{\theta}$, where $\vec{\theta}$ is the parameter vector and $\pi(\vec{\theta})$ represents the weighting or the deformation of the parameter space by the prior. Often, the logarithm $\log Z$ is computed. The Bayes factor $B_{12} = Z_1/Z_2$ is then compared to the *a priori* expectation. This method does not require models to be nested nor does it make assumptions about the parameter space or the data.

Alternatively, information criteria can be used. These are approximations in the limit which are valid given certain assumptions. The information criteria select a preferred model by comparing the difference in the best fit C -statistic between the models. Typically, the more complex model is punished by the additional number of parameters.

The Bayesian information criterion (BIC, Schwarz 1978) is an approximation to Bayesian model comparison. Unlike in the Bayesian evidence integration, only the maximum likelihood is needed, as the posterior is assumed to be strongly single-peaked, making the prior unimportant. Because the BIC uses the Laplace integral approximation, its results are in principle unreliable at the boundaries of the parameter space (e.g. checking whether a non-negative parameter is zero). The BIC decides, just like Bayesian model selection, which model has a higher probability of producing the data. The model with the smallest $BIC = C - m \times \ln n$ should be preferred, where n is the number of observations (data points) and m denotes the number of free parameters of the model.

The Akaike Information Criterion (AIC, Akaike 1974) is not rooted in Bayesian inference, but information theory. The AIC measures the information loss by using a specific model. Thus it can be said to consider the case of having multiple data sets and predicting the next. Technically, the Kullback-Leibler divergence, sometimes referred to as the “surprise” or “information gain” is minimised. The model with the smallest $AIC = C - 2 \times m$ should be preferred.

The robustness and efficiency of model comparison methods is tested via simulations in Buchner et al. (2014). There, it is shown that Bayesian model selection is more efficient and less prone to false selection than any method based on likelihood ratios (AIC, BIC, F-test and others).

5.2.3 Model verification

Any model comparison, or more generally any inference problem, assumes that one of the stated hypotheses is the true one. If no model is correct, the least bad model will be preferred. However, this assumption also requires examination.

Traditionally, this is quantitatively addressed by Goodness of Fit (GoF) measures such as the reduced χ^2 for normal distributions, with posterior-based approaches such as posterior predictive tests recently being developed (see e.g. Bayarri & Castellanos 2008). No consensus has been reached on these methods. Asymptotic results make assumptions not applicable in practice (e.g. the Likelihood-ratio and F-test are invalid at boundaries, making them unsuitable for feature-detection problems, see Protassov et al. 2002). Posterior-based approaches are diverse, and can be overly conservative (see e.g. Sinharay & Stern 2003). In the strict sense, there is no probabilistic framework available, frequentist or Bayesian, for testing whether the fitted model is correct. The fundamental reason for this is that outliers, although rare, are not strictly excluded like in classical logic. Additionally, a number of models can predict the same outcome.

However, the question can be relaxed to ask whether the model “explains” the data by looking at the range of values the observable typically covers (if the model is true). Monte Carlo simulations (parametric bootstrap) can be performed with the best-fit parameters

to check whether the spectrum can actually be produced by the given model. This can be computationally expensive however. It is easier to use a measure of “distance” between model and data (e.g. using the Kolmogorov-Smirnov statistic) to discover potentially problematic cases and to visually examine the data.

Q-Q plots (Wilk & Gnanadesikan 1968, see Section 4.4.2) provide a generic tool to visualise the goodness of fit, and are independent of the underlying distribution. The quantiles of the integrated observed counts are plotted against the integrated expected counts from the model. A good fit shows a straight line ($y = x$). This method is applied in Section 4.4 to the modelling of the background spectrum: The Q-Q plot is used for model discovery for improved fits, while model comparison using the AIC tests the significance of the model alterations. There, the AIC is appropriate, because a model is sought that retains the most information, and the models considered are not physical, so they are probably not the ones producing the data.

5.2.4 Priors

A model definition in the Bayesian framework would be incomplete without describing the priors for the models defined above. These encode an *a priori* weighting within the parameter space. Equivalently, a transformation of the unit space to the parameter space is sought. The simplest case is a uniform prior, giving equal weight in the parameter space. However, a uniform prior means something different whether the parameter or its logarithm is used. The parameters are very similar across the models used, so I describe them all together. They are also summarised in Table 5.1.

In the problem at hand, the prior distributions are assumed to be independent. Log-uniform priors are adopted for scale variables such as normalisations and N_H . The normalisation of spectral models is allowed to vary between 10^{-30} and 1. This conservatively contains all physically possible photon fluxes. However, since all considered models contain this parameter, the limits are irrelevant in practise. The normalisation of individual spectral components is then modelled relative to this model normalisation. N_H is defined from 10^{20} to 10^{26} cm^{-2} . The photon index Γ is modelled after the local sample analysed in Nandra & Pounds (1994), as a normal distribution with mean 1.95 and a standard deviation of 0.15. This encodes the assumption that distant AGNs behave like local AGNs with regard to their intrinsic spectrum. There is a degeneracy between N_H and Γ in that a steeper power law can be flattened by absorption. Hence, placing a prior on Γ plays an important role in constraining N_H . Nevertheless, I found that the results (e.g. the N_H distribution of the analysed sample) are relatively insensitive to other choices (for instance $\Gamma = 1.68 \pm 0.3$ from de Rosa et al. 2012), as the data drive the result in the majority of cases.

Of course we are not interested in our prior considerations, but in how the data strengthens or weakens hypotheses and re-weights parameter space. If the data have no discriminatory power, there is no information gain and the prior and posterior probability distributions will look the same. The difference between the prior and posterior can be measured using the Kullback-Leibler divergence (Kullback & Leibler 1951) as

$KL = \int posterior(x) \log \frac{posterior(x)}{prior(x)} dx$, essentially an integral difference across the parameter space. The KL divergence is measured in ban, a unit of information or entropy. A considerable information gain is e.g. $KL > 0.13$ bans which corresponds to halving the standard error of a Gaussian. I will restrict myself to computing the information gain only for the N_H parameter.

For model comparison, it is furthermore necessary to specify the *a priori* preference of models. Here, I consider two approaches: a) pair-wise comparisons from low to high complexity (e.g. `torus` vs. `torus+scattering`, see Figure 5.1.3), and between models of the same complexity. The scale of Jeffreys (1998) is adopted: A Bayes factor above 100 is “decisive”, 30 “very strong evidence”, 10-30 “strong evidence”, 3-10 “substantial evidence”. In $\log Z$, this corresponds to differences of 2, 1.5, 1 and 0.5 respectively. In case of a Bayes factor below 10 I remain with the simpler model. b) comparing all models simultaneously to find the model with the highest evidence. Both cases consider the models *a priori* equally probable.

5.2.5 Implementation

For the spectral analysis in the framework of Bayesian analysis, the `MultiNest` library was used and two software packages were created: `PyMultiNest` is a generic package for connecting Python probability functions with the `MultiNest` library, as well as model comparison and parameter estimation analysis of the `MultiNest` output. `BXA` is a package that connects the `Sherpa` X-ray analysis framework (Freeman et al. (2001)) with the Bayesian methodology presented here. `(Py)MultiNest` repeatedly suggests parameters on a unit hypercube which are transformed by `BXA` into model parameters using the prior definitions. `BXA` then computes a probability using `Sherpa`’s C-stat implementation, which is passed back to `(Py)MultiNest`. `(Py)MultiNest` can then be used to compute Bayes factors, create one or two-dimensional marginalised posterior probability distributions (PDFs) and output summarising Gaussian approximations.

`BXA` and `PyMultiNest` are publicly available on <http://github.com/JohannesBuchner>. In this work, `MultiNest` v2.17 is being used by `BXA` on `Sherpa` version 4.4v2 with 400 live points and a log-evidence accuracy of 0.1. For further analyses, I made extensive use of the `NumPy`/`SciPy`, `Matplotlib` and `Cosmology` packages (Jones et al. (2001–); Hunter (2007)<http://roban.github.com/CosmoPy/>) packages.

5.3 Results

The Bayesian methodology of parameter estimation and model comparison presented in Section 5.2 is applied using the models introduced in Section 5.1 to all sample spectra (CDFS only, see Section 4.9 for the selection). I first highlight the importance of the improved methodology, which is presented in Section 5.3.1. To demonstrate the model comparison, a single source is presented in detail in Section 5.3.3. Then model comparison is applied across the full sample (Section 5.3.2).

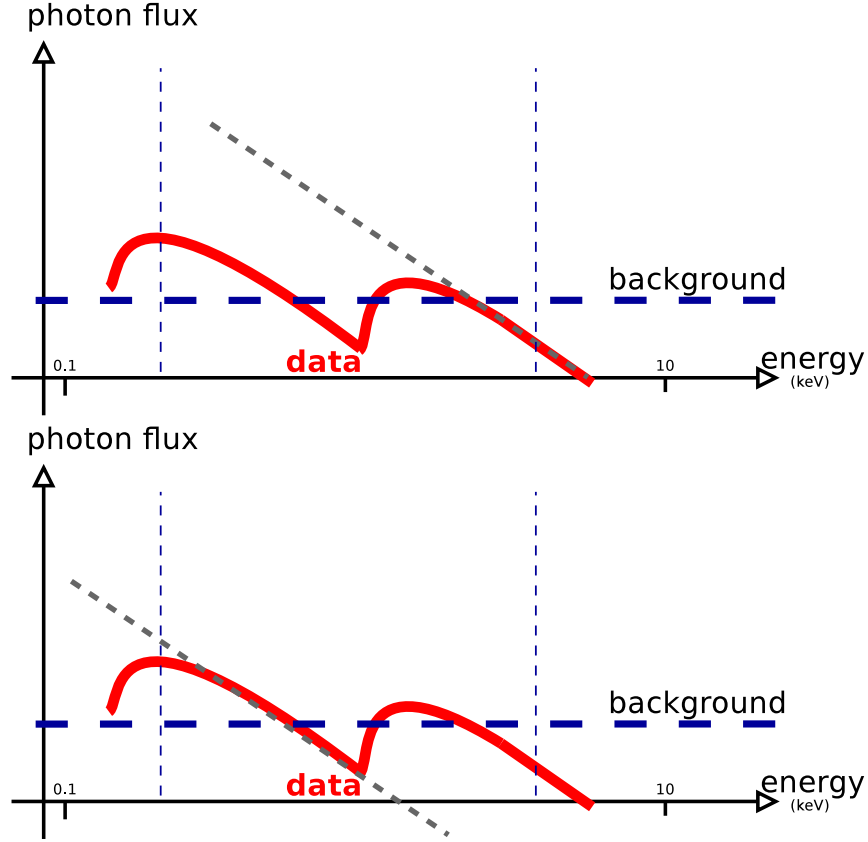


Figure 5.3.1: The origin of the two distinct solutions in Figure 5.3.2 is highlighted in these two cartoons. Given the data shown in the red thick line, one can either A) consider a bright, highly obscured source (top panel), or B) a low-luminosity, low-obscuration solution where the hard counts are due to the background. An intermediate solution is ruled out however. Depending on the background level and redshift, the two solutions will have different likelihoods.

5.3.1 Propagation of redshift uncertainty

It is common practise to use a redshift point estimator (best fit redshift) from photometric redshifts and analyse spectra with this value. In this work, in contrast, the redshift uncertainty from photometric redshift methods is propagated into the analysis of X-ray spectra in the form of a probability distribution on the redshift parameter. In this section, I discuss the differences between the approaches.

I demonstrate the strengths of the methodology described in the previous section using a particular source (ID 551 in the CDFS catalogue, RA=03h 32m 19.4s, Dec=-27° 42' 04"). The X-ray spectrum is analysed as laid out above using the `torus` model. The analysis is performed twice, (a) once with the probability distribution from photometric redshifts and (b) once with a fixed redshift value, namely the best fit photometric redshift. Figure 5.3.2 demonstrates how the different input redshifts (upper right panel) influence the results in

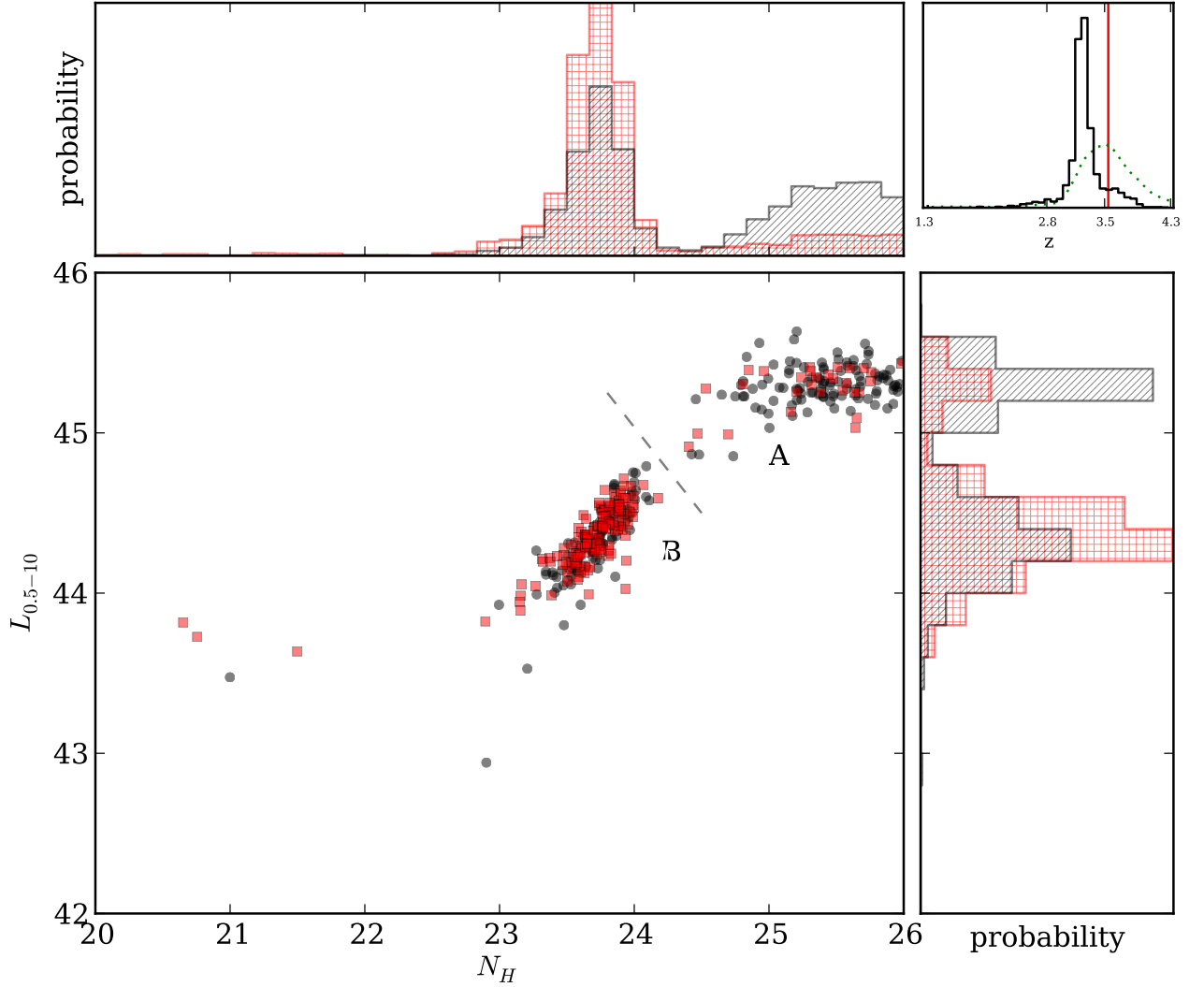


Figure 5.3.2: Demonstration of parameter estimation results under different redshift inputs. Source 551 in the CDFS catalogue is analysed twice with the methodology laid out in this paper using the `torus` model. I consider once the case of using only the best fit photometric redshift and once the case of using the photometric redshift probability distribution (PDZ). Both inputs are shown in the upper right panel as a vertical red line at $z = 3.5444$ and with a dotted green line respectively. For brevity, only two resulting parameters are presented, namely the column density N_H (logarithmic, in cm^{-2}) and the derived intrinsic luminosity L in the $0.5 - 10$ keV band (logarithmic, in erg/s). The large, lower left panel shows the derived intrinsic luminosity and column density parameters by equally probable points, similar to a Markov Chain. As each point on the plane is equally likely to be the true value, denser regions represent more probable parameter values. Here, the red squares represent the fixed redshift analysis while the black circles show the analysis results using the PDZ. The marginal distributions are shown in the upper left and lower right panels as probability histograms (red crossed hatching and black striped hatching). Two separated solutions are clearly visible and highlighted using the labels “A” for the highly obscured solution and “B” for the less obscured solution. The associated spectra are illustrated in Figure 5.3.1.

the derived column density and intrinsic luminosity parameters (large panel). The Bayesian analysis shows that the parameter space is broad, and split into two distinct solutions (see lower left panel in Figure 5.3.2): a highly obscured solution and a less obscured solution. The maximum likelihood is in the mildly obscured solution for the fixed redshift value, but in the heavily-obscured solution if the redshift distribution is used, because the likelihood improves when using a slightly lower value than the best fit redshift. The two solutions are strictly separated, i.e. an intermediate solution is ruled out. Commonly used Maximum Likelihood methods, like fitting and error estimation by Fisher matrix or contour search, will fail to estimate the uncertainty correctly and miss one of the two solutions, depending on the initial condition of the fit and the minimisation method adopted. Methods building on these results can therefore make false conclusions about e.g. the number of Compton-thick sources. The Bayesian inference method presented in this work can handle the separated solutions well and re-weights them based on the redshift information given. The difference in results comes from the freedom to move to a lower redshift, as can be seen in the upper right panel, where the black line shows the redshift posterior probability distribution. These results thus also show that redshift information can be improved using X-ray data (see Section 4.7.2).

5.3.2 Source 179: An example of Bayesian model comparison

Source 179 in the CDFS catalogue (spectroscopic $z = 0.605$, GOODS-MUSIC 15626) was detected with 2485 counts in the 0.5–10keV band at RA/Dec=(3h 32m 13.23s, -27°42'41.02"). This source was chosen because it illustrates the model selection well, showing several features, namely the Fe-K α line, scattering and absorption. In the next few paragraphs I present the source spectrum and how well different models reproduce them. Figure 5.3.3 overlays different models to the observed data. For brevity, only a subset of models are included in this presentation, namely **powerlaw**, **wabs**, **torus+scattering** and **wabs+pexmon+scattering**. I then show the results of model selection for this object in Table 5.2, where the log Z -column (normalised to highest) shows the computed evidence. Finally, the derived posterior parameters are shown.

A power law model (**powerlaw**) does not provide a good fit. This can be seen in the deviations between model and data points in the upper left panel, and also in the fact that this model has the lowest evidence of all models Table 5.2 (column 5). Furthermore, the derived photon index, $\Gamma = 0.8 \pm 0.05$ is unlikely and would constitute a 7σ outlier on the Gaussian distribution of photon indices of AGN (see Section 3.2.2).

Obscuration is expected in some sources, and the **wabs** model indeed improves the fit. The model follows the spectrum more closely (upper right panel in Figure 5.3.3) with a line-of-sight absorption of $N_H = 10^{22.5 \pm 0.1} \text{cm}^{-2}$. The evidence for this model is significantly higher, and rules out the **powerlaw** model. Here a difference of $\log Z_1 - \log Z_2 > \log 10 = 1$ is considered as a significant preference (see Section 5.2.4). Similarly, the **sphere** and **torus** models, although not shown here, improve the fit and are significantly preferred over the **powerlaw** model.

Comparing the data with the model prediction in the **wabs** spectrum, a line is visible

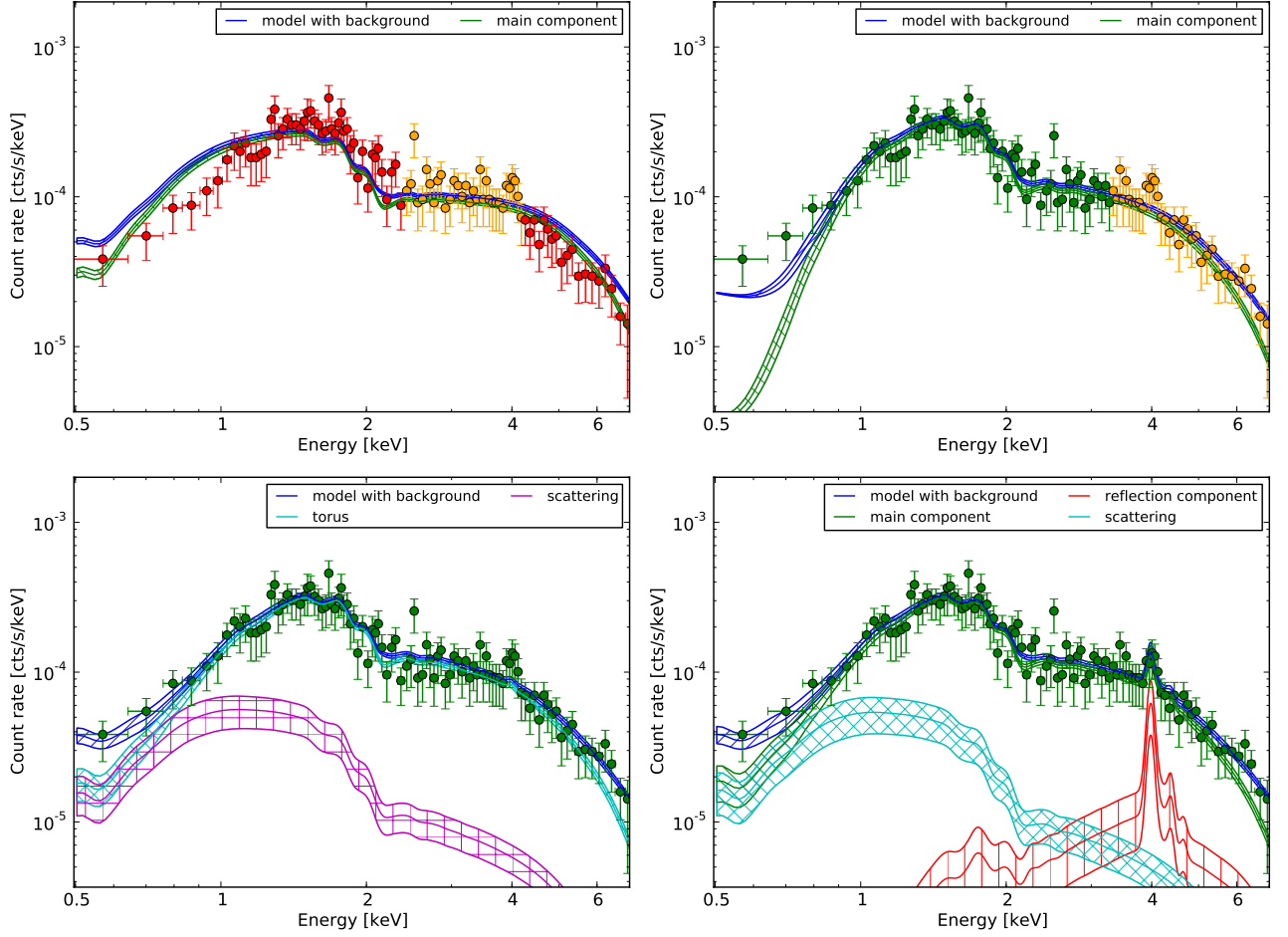


Figure 5.3.3: Observed spectrum and convolved models for object 179, binned for plotting to 10 counts per bin. Shown are analyses using various models and their individual components: **powerlaw** (upper left), **wabs** (upper right), **torus+scattering** (lower left) and **wabs+pexmon+scattering** (lower right). The posterior of the parameters is used to compute the median and 10%-quantiles of each model component.

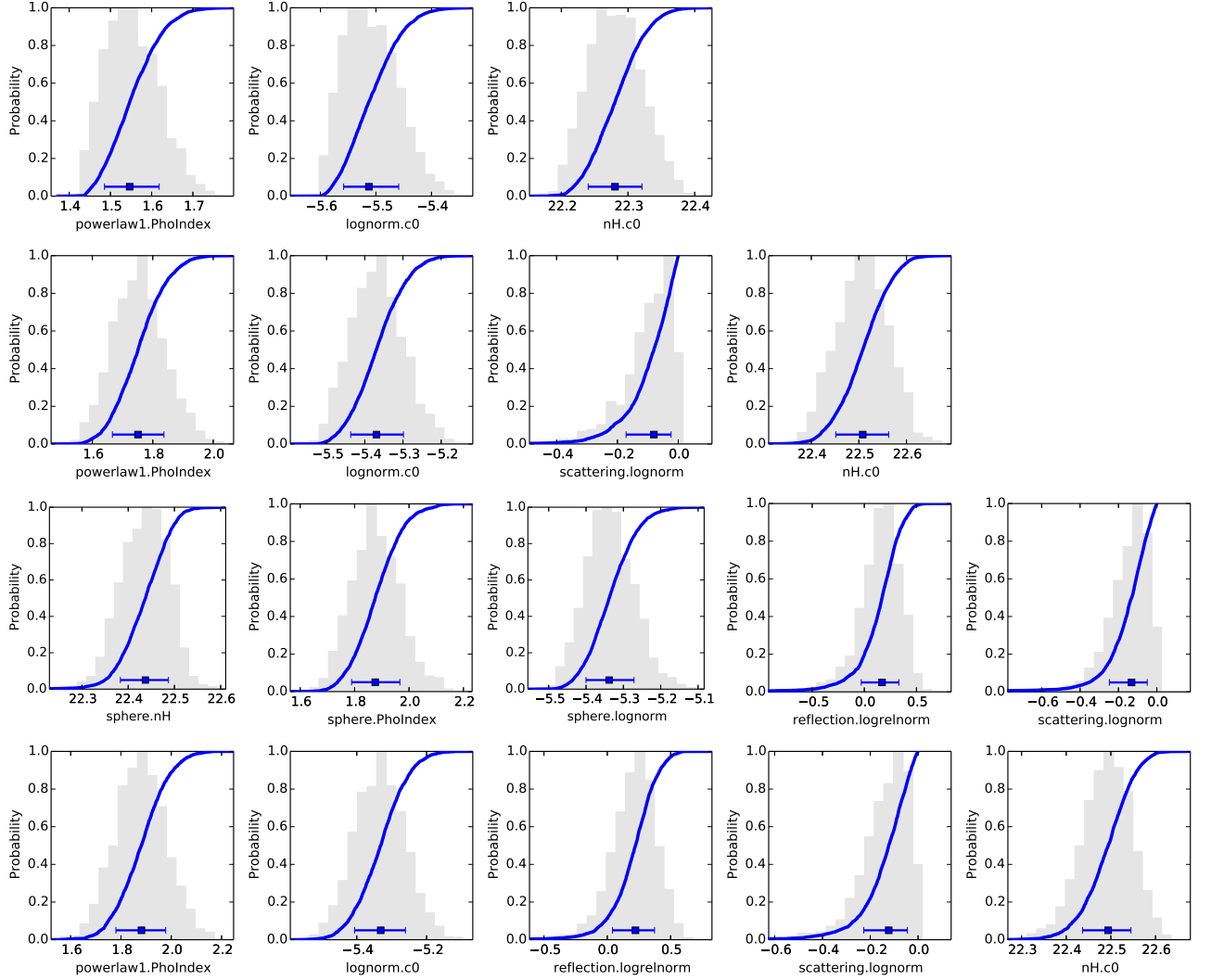


Figure 5.3.4: Marginalised parameters of the **wabs** (top), **wabs+scattering**, **sphere+pexmon+scattering** and **wabs+pexmon+scattering** model (bottom) for source 179. The posterior probability density distribution, normalised to the maximum, is shown by grey bars. The blue line indicates the cumulative posterior distribution. For summary of the error, the median and 10/90% quantiles can be used, or as the blue error bar indicates, the 1 standard-deviation equivalent probabilities.

Table 5.2: Model selection results for object 179. The columns indicate for each (1) considered model the derived (2) luminosity (logarithmic, in erg/s), (3) photon index, (4) column density (logarithmic, in cm^{-2}) with 1σ -equivalent quantile errors (see Figure 5.3.4 for the posterior distributions). The fifth column shows the information gain in column density in ban (see text). The remaining columns show the Bayesian model comparison results based on (6) the log-evidence Z (higher is more probable) and the (7) posterior probability for the given model, assuming all models have same prior probability. Additionally, the (8) BIC and (9) AIC statistics are shown as alternative model comparison techniques (lower is better).

Model	$L_{2-10\text{keV}}$	Γ	N_H	$\text{KL} _{N_H}$	$\log Z$	$p(M D)$	BIC	AIC
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
torus+pexmon+scattering	$43.35^{+0.01}_{-0.01}$	$1.87^{+0.1}_{-0.1}$	$22.44^{+0.05}_{-0.06}$	1.43	0.0	40.0%	-109.40	-122.59
sphere+pexmon+scattering	$43.35^{+0.01}_{-0.01}$	$1.88^{+0.1}_{-0.1}$	$22.44^{+0.05}_{-0.05}$	1.46	-0.1	30.2%	-109.40	-122.59
wabs+pexmon+scattering	$43.34^{+0.01}_{-0.01}$	$1.88^{+0.1}_{-0.1}$	$22.49^{+0.05}_{-0.06}$	1.44	-0.1	29.0%	-109.43	-122.62
torus+scattering	$43.39^{+0.01}_{-0.01}$	$1.76^{+0.1}_{-0.1}$	$22.45^{+0.05}_{-0.06}$	1.42	-2.0	0.4%	-105.18	-113.97
sphere+scattering	$43.39^{+0.01}_{-0.01}$	$1.77^{+0.1}_{-0.1}$	$22.46^{+0.05}_{-0.06}$	1.45	-2.1	0.3%	-105.87	-114.67
wabs+scattering	$43.39^{+0.01}_{-0.01}$	$1.75^{+0.1}_{-0.1}$	$22.51^{+0.05}_{-0.06}$	1.44	-2.7	0.1%	-103.08	-111.88
sphere	$43.35^{+0.01}_{-0.01}$	$1.57^{+0.1}_{-0.1}$	$22.23^{+0.05}_{-0.04}$	1.52	-4.7	0.0%	-98.07	-102.47
torus	$43.35^{+0.01}_{-0.01}$	$1.56^{+0.1}_{-0.1}$	$22.22^{+0.05}_{-0.04}$	1.54	-4.7	0.0%	-97.63	-102.02
wabs	$43.35^{+0.01}_{-0.01}$	$1.55^{+0.1}_{-0.1}$	$22.28^{+0.04}_{-0.04}$	1.58	-5.5	0.0%	-95.13	-99.53
powerlaw	$43.35^{+0.01}_{-0.01}$	$0.80^{+0.0}_{-0.0}$			-27.9	0.0%	0.00	0.00

at $\sim 4\text{keV}$ as well as an excess of soft energy counts. The former coincides with the Fe-K α line, while the latter may be modeled with the `+scattering` component. Considering `torus+scattering` and `wabs+pexmon+scattering` (lower panels in Figure 5.3.3), they both model the observed counts well. However, the Fe-K α line is clearly visible in the data. Compared to the simple absorption models, the addition of `+scattering` and `+pexmon` increases the evidence (see Table 5.2), hence these models are preferred, while the other models, e.g. the widely used `wabs+scattering`, are ruled out for this source.

When comparing models, an important aspect is the model-dependence of derived physical parameters. Table 5.2 shows the intrinsic luminosity, photon index and line-of-sight column density. These are computed by the posterior values and summarised using 1σ -equivalent quantiles. The marginal posterior distributions are shown in Figure 5.3.4. Simple absorption model like `wabs` try to compensate for the lack of soft X-rays by using a flatter spectrum with less absorption. Both the `+scattering` and `+pexmon` component increase the photon index. As these additional components only take up a fraction of the intrinsic power law component, the changes in intrinsic luminosity are small. Despite these changes, the derived values for the luminosity and obscuring column are consistent between obscurer geometries (i.e. `wabs`, `torus`, `sphere`).

Furthermore, it is important to check whether the derived physical parameters are strongly influenced by the prior. For example, for a weak source with data of no discriminatory power, the posterior of N_H would look like the prior (log-uniform). The KL-divergence (see Section 5.2.4), also known as the information gain or knowledge update, is used to measure the “distance” between prior and posterior of N_H . For this particular source, the values are shown in column 4 of Table 5.2. For reference, a considerable information gain is $KL > 0.13$ bans, which corresponds to narrowing the width of a Gaussian distribution by a factor of two ($\sigma_{\text{Posterior}} = \sigma_{\text{Prior}}/2$). Because N_H is well-constrained in this source compared to the prior, the $KL|_{N_H}$ values are high, and the posterior is not dominated by the prior.

The BIC and AIC values (column 7 and 8 of Table 5.2, lower values are preferred) show the same preferences as the Bayesian evidence computations. They are an approximate method of model selection based on the likelihood ratio and parameter penalisation. There are however important differences to the Bayesian model selection (see Section 5.2.2 for details). The evidence and its approximation, the BIC, can be used to express how much more probable a model is compared to the others based on the data (column 6). The AIC, in contrast, measures the information loss by using a specific model. In this single source, the model selection prefers models with absorption, scattering and an additional reflection component (`pexmon`). However, no preference is found between the geometries of `wabs`, `torus` and `sphere`. To improve the discriminatory power, the evidence of the full sample is combined.

5.3.3 Model selection on the full sample

This section presents the results for the full sample in two forms. Firstly, in each source, model comparison is performed between each pair of models (arrows in Figure 5.1.3). I

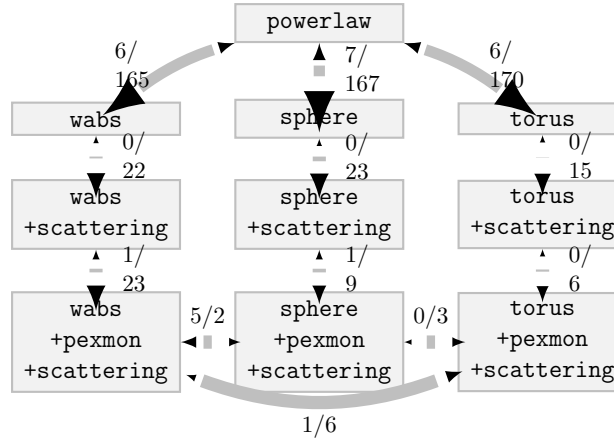


Figure 5.3.5: Model comparison preferences of the sample. The arrow size and numbers indicate the number of sources for which one model is strongly preferred over the other.

count for how many sources a preference was found. This is shown in Figure 5.3.5, where the two numbers indicate the two directions. For instance, in 6 sources, the (simpler) **powerlaw** model is preferred over the **torus**, while in 170 sources, the **torus** is preferred. The size of the arrows visualises the same information. There is clear preference for absorption, scattering and reflection in ~ 170 , ~ 20 and > 5 cases respectively, showing that in a substantial number of sources these components are required. Between **wabs+pexmon+scattering**, **torus+pexmon+scattering** and **sphere+pexmon+scattering**, there is no clear trend, however the **torus** geometry is preferred most often. This indicates some variety in the obscurer geometry between sources.

To distinguish between the remaining three models, stronger data constraints are necessary. These can come from combining a sample of AGN, assuming one of the models describes all objects, albeit with different parameters. Mathematically, this is done by multiplying the evidence values Z to obtain the evidence for each model across all datasets. This is shown in Table 5.3. For each model, the evidence is stacked in column 3 by summing the $\log Z$ values across the sample of 334 sources. For comparison, column 2 shows the number of sources for which the model was ruled out by the other models. As shown with Source 179 above, individual sources can already have strong preferences in the model selection. The sum would then be dominated by the contribution of a single object. To have a result descriptive of the sample that is not dominated by outliers, I apply bootstrapping. Sources are drawn with repetition from the sample, and the same quantities (number of rejections, summed $\log Z$ values) are computed, and taking all draws together the mean and root mean square is estimated. Additionally, for each draw it is computed whether this model is ruled out ($\sum \log Z_1 > \sum \log Z_2 + 1$). If the model is ruled out in 100% of draws (column 6), the result is robust against bootstrapping and thus there is confidence that this result will also hold for the parent sample and that is not dominated by outliers.

The power law model (**powerlaw**) has the lowest evidence and is ruled out by simple

Table 5.3: Sample model comparison. Considering each source in turn, the second column shows in how many sources model comparison ruled out the particular model. The third column shows the total evidence across all sources, relative to the best evidence. The following columns show the same statistics, but using bootstrapping on the sample, making the computed quantities more robust against outliers. The last column computes how often in the bootstrapping the model was ruled out based on the total evidence (10 times less likely than other models).

Model (1)	Sample results		Bootstrapped results		
	# rej (2)	$\sum \log Z$ (3)	# rej (4)	$\sum \log Z$ (5)	ruled out (6)
<i>All (334 sources)</i>					
torus+pexmon+scattering	11	-85.6	10 ± 3.4	-85.4 ± 8.0	$9 \pm 28\%$
wabs+pexmon+scattering	8	-93.4	8 ± 2.7	-93.6 ± 7.6	$85 \pm 35\%$
sphere+pexmon+scattering	11	-95.3	10 ± 3.3	-95.1 ± 7.8	$100 \pm 5\%$
torus+scattering	16	-116.8	16 ± 4.0	-117.3 ± 11.8	$100 \pm 0\%$
sphere+scattering	21	-127.3	21 ± 4.6	-127.3 ± 10.3	$100 \pm 0\%$
wabs+scattering	37	-167.6	37 ± 5.7	-168.6 ± 14.2	$100 \pm 0\%$
torus	39	-188.2	39 ± 5.9	-190.4 ± 20.0	$100 \pm 0\%$
sphere	55	-233.1	55 ± 6.8	-235.3 ± 22.3	$100 \pm 0\%$
wabs	59	-270.9	59 ± 6.8	-274.1 ± 26.9	$100 \pm 0\%$
powerlaw	187	-2657.9	186 ± 9.6	-2704.8 ± 412.6	$100 \pm 0\%$

absorption models. These in turn are ruled out by absorption with additional scattering (in Figure 5.3.5, strongly preferred in 15-23 objects). Then, for 6-23 objects additional **pexmon** reflection is strongly preferred.

The remaining models are thus such that absorption, scattering and reflection are required. Between **wabs+pexmon+scattering**, **sphere+pexmon+scattering** and **torus+pexmon+scattering**, the number of objects for which a model is rejected remains comparable. The results for the full sample have large and significant differences in $\sum \log Z$ in Table 5.3, with the **torus+pexmon+scattering** model having the highest evidence. But when bootstrapping (column 5 and 6 in Table 5.3), the values show large variation and overlap broadly, showing that the difference is not robust against bootstrapping, and sample-dependent. This indicates considerable variation in evidence, i.e. variation in the models preferred, between sources. This is investigated by splitting the samples by N_H , shown in Table 5.4. For this, a source is assigned to a N_H bin if both the 10% and 90% quantiles of the posterior values determined from **torus+scattering** fall inside. As indicated before, N_H is comparable between absorption models. In the Compton-thick regime ($N_H \gtrsim 10^{24} \text{cm}^{-2}$), **torus+pexmon+scattering** is consistently strongly preferred over e.g. **wabs+pexmon+scattering** (forth segment in Table 5.4). However, the scatter in the bootstrapped results for the full sample (Table 5.3) indicates that some sources must favour **wabs+pexmon+scattering**, while others favour **torus+pexmon+scattering**. In Figure 5.3.7, the dependence of the preference is shown as a function of obscuring column.

Table 5.4: Sub-sample model comparison. As in Table 5.3, but the sample is split by $\log N_H$, estimated using the 90% quantiles of the **torus+scattering** model posterior, or redshift. Models with mean 100% and root mean square 0% in column 6 are not shown in the lower segments.

Model (1)	Sample results		Bootstrapped results		
	# rej (2)	$\sum \log Z$ (3)	# rej (4)	$\sum \log Z$ (5)	ruled out (6)
$\log N_H = 20 - 22$ (54 sources)					
wabs+pexmon+scattering	2	-24.0	2 ± 1.3	-24.4 ± 5.1	$54 \pm 50\%$
sphere+pexmon+scattering	2	-24.4	2 ± 1.3	-24.9 ± 5.3	$63 \pm 48\%$
powerlaw	7	-24.9	7 ± 2.7	-25.2 ± 9.4	$44 \pm 50\%$
$\log N_H = 22 - 23$ (47 sources)					
torus+pexmon+scattering	0	-5.3	0 ± 0.0	-5.2 ± 0.7	$0 \pm 0\%$
sphere+pexmon+scattering	0	-6.6	0 ± 0.0	-6.5 ± 0.7	$70 \pm 46\%$
wabs+pexmon+scattering	0	-7.5	0 ± 0.0	-7.5 ± 0.7	$98 \pm 14\%$
torus+scattering	3	-13.1	2 ± 1.6	-12.6 ± 3.5	$98 \pm 14\%$
$\log N_H = 23 - 24$ (51 sources)					
wabs+pexmon+scattering	0	-10.1	0 ± 0.0	-10.1 ± 1.1	$6 \pm 24\%$
torus+pexmon+scattering	4	-12.6	3 ± 1.8	-12.4 ± 2.2	$67 \pm 47\%$
$\log N_H = 24 - 26$ (14 sources)					
torus+scattering	0	-1.5	0 ± 0.0	-1.5 ± 0.5	$7 \pm 26\%$
torus+pexmon+scattering	0	-1.7	0 ± 0.0	-1.7 ± 0.3	$7 \pm 26\%$
torus	3	-9.8	3 ± 2.1	-10.3 ± 7.1	$93 \pm 26\%$
$\log N_H = 22 - 26$ (176 sources)					
torus+pexmon+scattering	4	-30.2	4 ± 1.9	-30.0 ± 3.1	$0 \pm 0\%$
wabs+pexmon+scattering	6	-46.4	6 ± 2.5	-46.6 ± 5.1	$99 \pm 8\%$
$z > 1$ (229 sources)					
torus+pexmon+scattering	8	-57.6	8 ± 3.0	-58.9 ± 7.7	$11 \pm 32\%$
wabs+pexmon+scattering	5	-64.5	5 ± 2.3	-65.2 ± 7.7	$80 \pm 40\%$
torus+scattering	7	-67.7	7 ± 2.5	-68.5 ± 7.7	$100 \pm 7\%$
$z > 2$ (96 sources)					
torus+pexmon+scattering	4	-28.4	4 ± 2.0	-28.8 ± 6.1	$25 \pm 43\%$
torus+scattering	3	-30.2	2 ± 1.6	-30.8 ± 6.0	$75 \pm 43\%$
wabs+pexmon+scattering	2	-30.8	2 ± 1.3	-31.7 ± 5.9	$66 \pm 47\%$
sphere+scattering	7	-39.0	7 ± 2.4	-39.9 ± 5.8	$99 \pm 10\%$

Only highly obscured sources $N_H > 10^{23} \text{cm}^{-2}$ can significantly decide between models. In the Compton-thick regime, `torus+pexmon+scattering` is preferred, in a number of sources. However, just below the Compton-thick regime, `sphere+pexmon+scattering` is preferred.

A further concern might be that low-redshift sources with many counts dominate the result, ignoring the main target population of our interest. The last two segments of Table 5.4 show the result of selecting only sources with $z > 1$ and $z > 2$ respectively. The inference results in this regime are entirely consistent with the results for the full sample.

Overall however, `torus+pexmon+scattering` can be considered the best model. A catalogue of the derived quantities for each source in the CDFS was compiled using this model (Table 5.5 shows an excerpt, Table 5.6 lists all Compton-thick sources; the complete catalogue is available online in Buchner et al. 2014). These tables list column densities, intrinsic power law index, intrinsic luminosity as well as the relative normalisations of the additional scattering and reflection components. The most important parameters for e.g. luminosity function studies are $L_{2-10\text{keV}}$, z and N_H , which are visualised in Figure 5.3.6.

Figure 5.3.9 shows a comparison to previously published works in the CDFS, specifically that of Tozzi et al. (2006). There are important differences between the works in terms of methodology and data. The Tozzi et al. (2006) analysis is based on only the first 1Ms of CDFS data, and thus has many fewer counts. Furthermore, only simple absorption models have been considered in their maximum likelihood fitting. Except for low count sources with large uncertainties, there is overall agreement between the results. The found Compton-thick AGN (Table 5.6) are in agreement with the sample found by Brightman & Ueda (2012), except that the selection criteria used in this work remove a number of sources whose soft photons may be dominated by stellar processes. One source (ID 186 in their paper), is not found to be a Compton-thick AGN, as a different redshift from the improved catalogue (Hsu et al. 2014) was used.

5.4 Discussion

5.4.1 X-ray spectral analysis methodology

Before putting the results into context, I review the methodology.

I have presented a new framework and method for analysing X-ray spectra, relying on Bayesian inference using nested sampling. In particular, parameter estimation and model comparison are easily possible and overcome considerable limitations of current methods (see Section 5.2 for a detail discussion of various methods):

1. **No binning of data.** Low-count and high-count sources are treated the same way using Poisson statistics, as with C-stat in the well-established maximum likelihood estimation methods. No information loss from binning needs to be introduced.
2. **Background modelling.** The background is modelled with a continuous non-physical model (Gaussian mixture). Unlike other options (background subtractions,

Table 5.5: Catalogue (excerpt). The parameters are derived using the **torus+pexmon+scattering** model. (1) XID, (2,3) Right ascension and Declination (JD2000) in degrees, (4) Photon counts obtained in the analysed $0.5 - 8\text{keV}$ range. (5) Redshift with (posterior) uncertainties if photometric (shortened to 3 significant digits). (6) Intrinsic luminosity ($2-10\text{keV}$) in erg/s . (7) logarithm of the column density N_H in cm^{-2} . (8) Photon index; The prior was 1.95 ± 0.15 , so if no information was gained this value remains. (9) Scattering normalization relative to the intrinsic power law. (10) Reflection normalization of the **pexmon** component relative to the intrinsic power law. (11) Information gain measured from the N_H posterior in bans. As a reference, the narrowing of a Gaussian from prior to posterior by a factor of 2 corresponds to 0.13ban , and thus values higher than that correspond to significant discriminatory information in the data. (12) Annotations; S when $f_{\text{scat}} > 3\%$, s when $f_{\text{scat}} > 0.5\%$ with $\geq 90\%$ probability; R when $R > 0.3$ with $\geq 90\%$ probability, i.e. strong additional **pexmon** reflection; Compton-thick (CT) if $N_H > 10^{24}\text{cm}^{-2}$, Compton-thin (O) if $10^{22}\text{cm}^{-2} < N_H < 10^{24}\text{cm}^{-2}$, Unobscured (U, $N_H < 10^{22}\text{cm}^{-2}$, each with $\geq 50\%$ probability. Only sources with R, s or S are shown here.

ID	RA	Dec	counts	z	$L_{2-10\text{keV}}$	N_H	Γ	f_{scat}	R	$KL _{N_H}$	notes
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
49	3:32:14.01	-27:51:00.53	1997	0.122	$41.65^{+0.01}_{-0.01}$	$22.17^{+0.02}_{-0.02}$	$1.80^{+0.05}_{-0.05}$	$2.2^{+0.4\%}_{-0.4\%}$	$0.06^{+0.06}_{-0.03}$	1.49	O,s
242	3:31:52.35	-27:47:52.79	762	$1.84^{+0.02}_{-0.03}$	$43.04^{+0.03}_{-0.03}$	$21.62^{+0.31}_{-0.44}$	$1.87^{+0.05}_{-0.05}$	$0.0^{+0.0\%}_{-0.0\%}$	$6.31^{+1.30}_{-1.33}$	0.36	U,R
21	3:32:12.94	-27:52:36.57	957	2.562	$44.12^{+0.03}_{-0.03}$	$23.21^{+0.03}_{-0.03}$	$1.76^{+0.06}_{-0.04}$	$0.0^{+0.0\%}_{-0.0\%}$	$1.08^{+0.37}_{-0.31}$	1.34	O,R
131	3:32:22.54	-27:46:03.84	597	1.727	$43.75^{+0.05}_{-0.05}$	$23.37^{+0.04}_{-0.05}$	$1.85^{+0.05}_{-0.05}$	$5.9^{+0.8\%}_{-0.8\%}$	$1.25^{+0.66}_{-0.68}$	1.25	O,S
118	3:31:52.53	-27:46:42.30	1295	0.673	$42.63^{+0.02}_{-0.02}$	$21.77^{+0.06}_{-0.06}$	$1.90^{+0.07}_{-0.06}$	$0.0^{+0.0\%}_{-0.0\%}$	$2.18^{+0.64}_{-0.61}$	1.02	U,R
338	3:32:49.20	-27:40:50.01	667	$1.26^{+0.00}_{-0.00}$	$43.93^{+0.08}_{-0.08}$	$23.88^{+0.06}_{-0.06}$	$1.99^{+0.06}_{-0.06}$	$3.6^{+0.9\%}_{-0.7\%}$	$0.14^{+0.25}_{-0.08}$	1.05	O,s
144	3:32:29.98	-27:45:29.96	3961	1.218	$43.81^{+0.01}_{-0.01}$	$20.52^{+0.19}_{-0.17}$	$1.75^{+0.02}_{-0.02}$	$0.0^{+0.0\%}_{-0.0\%}$	$1.14^{+0.24}_{-0.21}$	0.71	U,R
104	3:32:08.66	-27:47:34.34	28019	0.543	$43.83^{+0.00}_{-0.00}$	$20.82^{+0.04}_{-0.05}$	$1.95^{+0.01}_{-0.01}$	$0.0^{+0.0\%}_{-0.0\%}$	$0.79^{+0.09}_{-0.09}$	1.21	U,R
190	3:32:27.00	-27:41:05.11	32099	0.734	$44.20^{+0.00}_{-0.00}$	$20.08^{+0.04}_{-0.03}$	$1.97^{+0.01}_{-0.01}$	$0.0^{+0.0\%}_{-0.0\%}$	$0.78^{+0.08}_{-0.08}$	1.36	U,R
92	3:31:58.11	-27:48:33.97	915	0.734	$42.53^{+0.02}_{-0.02}$	$21.26^{+0.14}_{-0.18}$	$1.84^{+0.07}_{-0.06}$	$0.0^{+0.0\%}_{-0.0\%}$	$3.39^{+0.84}_{-0.72}$	0.61	U,R
179	3:32:13.23	-27:42:41.02	2485	0.605	$43.17^{+0.01}_{-0.01}$	$22.45^{+0.02}_{-0.03}$	$1.90^{+0.05}_{-0.04}$	$7.4^{+0.8\%}_{-0.9\%}$	$1.61^{+0.29}_{-0.31}$	1.45	O,S,R
287	3:31:55.40	-27:54:47.21	618	0.737	$43.72^{+0.03}_{-0.03}$	$23.02^{+0.04}_{-0.04}$	$1.89^{+0.07}_{-0.06}$	$3.4^{+0.6\%}_{-0.6\%}$	$0.13^{+0.19}_{-0.08}$	1.31	O,s
4	3:32:38.90	-27:57:00.17	4687	0.297	$43.06^{+0.01}_{-0.01}$	$22.59^{+0.01}_{-0.01}$	$1.91^{+0.04}_{-0.04}$	$2.5^{+0.2\%}_{-0.2\%}$	$1.16^{+0.18}_{-0.18}$	1.76	O,s,R
133	3:32:39.08	-27:46:01.79	1355	1.216	$43.56^{+0.02}_{-0.02}$	$22.60^{+0.03}_{-0.03}$	$2.11^{+0.05}_{-0.05}$	$9.4^{+0.3\%}_{-0.4\%}$	$0.87^{+0.43}_{-0.43}$	1.26	O,S
50	3:32:18.34	-27:50:55.13	647	1.536	$44.18^{+0.04}_{-0.05}$	$23.80^{+0.03}_{-0.03}$	$1.89^{+0.06}_{-0.06}$	$1.0^{+0.2\%}_{-0.2\%}$	$0.17^{+0.24}_{-0.10}$	1.46	O,s
181	3:32:47.88	-27:42:32.78	8718	0.979	$44.12^{+0.00}_{-0.00}$	$22.30^{+0.02}_{-0.02}$	$1.86^{+0.02}_{-0.02}$	$7.3^{+1.9\%}_{-1.1\%}$	$0.05^{+0.04}_{-0.02}$	1.56	O,S
132	3:32:03.66	-27:46:03.74	1380	0.574	$43.07^{+0.02}_{-0.02}$	$22.79^{+0.02}_{-0.02}$	$1.88^{+0.06}_{-0.05}$	$0.5^{+0.4\%}_{-0.5\%}$	$1.26^{+0.32}_{-0.30}$	1.56	O,R
7	3:32:40.82	-27:55:46.76	1557	$3.25^{+0.01}_{-0.01}$	$44.86^{+0.04}_{-0.04}$	$23.91^{+0.02}_{-0.02}$	$1.86^{+0.06}_{-0.05}$	$3.6^{+0.6\%}_{-0.6\%}$	$0.06^{+0.06}_{-0.03}$	1.49	O,s
200	3:32:34.39	-27:39:13.55	1848	$1.56^{+0.01}_{-0.01}$	$44.46^{+0.02}_{-0.02}$	$23.09^{+0.02}_{-0.03}$	$2.01^{+0.04}_{-0.05}$	$4.8^{+0.7\%}_{-0.7\%}$	$0.20^{+0.26}_{-0.13}$	1.45	O,s

Table 5.6: Catalogue (Compton-thick AGN only). Refer to Table 5.5 for a description of the columns.

ID	RA	Dec	counts	z	$L_{2-10\text{keV}}$	N_H	Γ	f_{scat}	R	KL_{NH}	notes
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
460	3:32:23.41	-27:42:55.89	194	2.145	43.89 ^{+0.16} _{-0.19}	24.37 ^{+0.15} _{-0.13}	1.94 ^{+0.07} _{-0.06}	0.0 ^{+0.0%} _{-0.0%}	0.32 ^{+0.56} _{-0.23}	0.61	CT
420	3:32:23.60	-27:46:01.23	48	1.033	43.11 ^{+0.12} _{-0.15}	24.32 ^{+0.07} _{-0.09}	1.93 ^{+0.08} _{-0.07}	0.0 ^{+0.0%} _{-0.0%}	3.45 ^{+2.03} _{-1.81}	0.91	CT
412	3:32:42.87	-27:48:09.44	67	2.42 ^{+0.04} _{-0.05}	43.69 ^{+0.16} _{-0.19}	24.42 ^{+0.14} _{-0.13}	1.94 ^{+0.06} _{-0.06}	0.0 ^{+0.0%} _{-0.0%}	0.28 ^{+0.51} _{-0.20}	0.59	CT
448	3:32:19.41	-27:40:51.76	528	0.682	43.71 ^{+0.07} _{-0.09}	25.07 ^{+0.33} _{-0.32}	1.99 ^{+0.08} _{-0.07}	0.3 ^{+0.1%} _{-0.1%}	0.55 ^{+1.09} _{-0.40}	0.37	CT
474	3:32:15.81	-27:42:06.78	339	1.71 ^{+0.21} _{-0.05}	43.85 ^{+0.21} _{-0.22}	24.21 ^{+0.15} _{-0.15}	1.95 ^{+0.06} _{-0.06}	0.0 ^{+0.0%} _{-0.0%}	0.19 ^{+0.38} _{-0.12}	0.55	CT
271	3:32:33.86	-27:42:04.30	306	1.45 ^{+0.01} _{-0.01}	43.20 ^{+0.42} _{-0.99}	24.16 ^{+0.34} _{-2.29}	1.95 ^{+0.05} _{-0.06}	0.4 ^{+1.0%} _{-0.4%}	1.03 ^{+1.71} _{-0.69}	0.25	CT
478	3:32:14.60	-27:52:57.06	164	1.24 ^{+0.01} _{-0.01}	43.78 ^{+0.09} _{-0.13}	24.84 ^{+0.29} _{-0.24}	1.95 ^{+0.06} _{-0.06}	0.0 ^{+0.0%} _{-0.0%}	0.41 ^{+0.91} _{-0.27}	0.36	CT
467	3:32:00.80	-27:53:33.84	541	1.60 ^{+0.02} _{-0.03}	43.55 ^{+0.17} _{-0.17}	24.14 ^{+0.09} _{-0.08}	1.91 ^{+0.06} _{-0.06}	0.0 ^{+0.0%} _{-0.0%}	2.05 ^{+2.01} _{-1.40}	0.85	CT
437	3:32:39.15	-27:48:32.26	54	2.470	43.60 ^{+0.13} _{-0.15}	24.27 ^{+0.10} _{-0.09}	1.94 ^{+0.06} _{-0.06}	0.0 ^{+0.0%} _{-0.0%}	0.36 ^{+0.72} _{-0.26}	0.84	CT
345	3:32:24.67	-27:54:43.24	255	0.123	41.41 ^{+0.41} _{-1.21}	24.13 ^{+0.74} _{-2.02}	1.95 ^{+0.06} _{-0.07}	0.5 ^{+0.6%} _{-0.2%}	0.37 ^{+0.96} _{-0.28}	0.16	CT
451	3:32:31.49	-27:50:28.64	38	1.613	42.93 ^{+0.96} _{-0.96}	24.34 ^{+0.43} _{-0.55}	1.95 ^{+0.07} _{-0.07}	0.0 ^{+0.1%} _{-0.0%}	0.40 ^{+0.88} _{-0.30}	0.11	CT
266	3:32:14.94	-27:42:24.83	976	1.81 ^{+0.02} _{-0.02}	44.75 ^{+0.14} _{-1.38}	24.95 ^{+0.46} _{-3.94}	1.94 ^{+0.05} _{-0.05}	2.0 ^{+0.8%} _{-2.0%}	1.94 ^{+3.60} _{-1.65}	0.35	CT
416	3:32:21.77	-27:46:56.99	64	1.32 ^{+0.03} _{-0.02}	43.31 ^{+0.10} _{-0.12}	24.09 ^{+0.05} _{-0.05}	1.94 ^{+0.06} _{-0.07}	0.0 ^{+0.0%} _{-0.0%}	0.28 ^{+0.59} _{-0.18}	1.08	CT
406	3:31:45.16	-27:49:48.72	571	3.153	44.44 ^{+0.14} _{-0.24}	24.73 ^{+0.18} _{-0.18}	1.91 ^{+0.07} _{-0.05}	0.0 ^{+0.0%} _{-0.0%}	1.09 ^{+1.29} _{-0.75}	0.45	CT
400	3:32:25.18	-27:54:49.72	346	1.090	44.36 ^{+0.03} _{-0.04}	25.23 ^{+0.26} _{-0.37}	1.88 ^{+0.05} _{-0.04}	0.0 ^{+0.0%} _{-0.0%}	0.94 ^{+1.39} _{-0.61}	0.46	CT
404	3:32:36.15	-27:50:36.97	108	1.608	43.65 ^{+0.12} _{-0.12}	24.07 ^{+0.08} _{-0.08}	1.96 ^{+0.07} _{-0.07}	0.2 ^{+0.4%} _{-0.2%}	0.16 ^{+0.32} _{-0.10}	0.91	CT
401	3:31:50.45	-27:52:11.49	592	1.370	44.44 ^{+0.08} _{-0.25}	25.17 ^{+0.27} _{-0.45}	1.95 ^{+0.07} _{-0.07}	0.5 ^{+0.2%} _{-0.3%}	1.02 ^{+2.52} _{-0.81}	0.27	CT
158	3:32:22.59	-27:44:25.97	129	0.738	43.38 ^{+0.17} _{-0.20}	24.25 ^{+0.11} _{-0.10}	1.95 ^{+0.06} _{-0.06}	0.4 ^{+0.3%} _{-0.2%}	0.75 ^{+1.48} _{-0.54}	0.73	CT
77	3:32:35.71	-27:49:16.18	95	2.578	44.43 ^{+0.04} _{-0.14}	25.50 ^{+0.12} _{-0.14}	1.93 ^{+0.06} _{-0.06}	0.0 ^{+0.0%} _{-0.0%}	0.88 ^{+1.41} _{-0.68}	0.75	CT
417	3:32:21.97	-27:46:56.00	35	0.670	43.18 ^{+0.07} _{-0.09}	24.99 ^{+0.30} _{-0.40}	1.92 ^{+0.07} _{-0.07}	0.0 ^{+0.0%} _{-0.0%}	0.26 ^{+0.50} _{-0.19}	0.41	CT
570	3:33:00.76	-27:48:57.46	438	1.67 ^{+0.04} _{-0.08}	43.76 ^{+0.19} _{-0.18}	24.36 ^{+0.11} _{-0.12}	1.93 ^{+0.07} _{-0.06}	1.0 ^{+1.3%} _{-0.9%}	1.24 ^{+1.49} _{-0.89}	0.64	CT
226	3:32:14.42	-27:51:10.43	141	1.544	43.68 ^{+0.14} _{-0.13}	24.16 ^{+0.07} _{-0.06}	1.96 ^{+0.07} _{-0.07}	1.2 ^{+0.7%} _{-0.5%}	0.23 ^{+0.55} _{-0.14}	0.94	CT
441	3:32:22.79	-27:45:28.54	46	1.89 ^{+0.04} _{-0.04}	43.15 ^{+0.23} _{-0.47}	24.44 ^{+0.23} _{-0.24}	1.93 ^{+0.06} _{-0.06}	0.0 ^{+0.0%} _{-0.0%}	0.45 ^{+1.00} _{-0.33}	0.26	CT
408	3:32:33.22	-27:49:15.92	38	3.55 ^{+0.03} _{-0.03}	43.63 ^{+0.30} _{-0.24}	24.29 ^{+0.54} _{-0.22}	1.95 ^{+0.07} _{-0.06}	0.0 ^{+0.0%} _{-0.0%}	0.26 ^{+0.40} _{-0.16}	0.40	CT
459	3:32:29.48	-27:43:22.27	143	1.609	44.07 ^{+0.05} _{-0.06}	25.01 ^{+0.14} _{-0.13}	1.90 ^{+0.06} _{-0.05}	0.0 ^{+0.0%} _{-0.0%}	1.77 ^{+1.81} _{-1.21}	0.60	CT
384	3:32:42.02	-27:39:49.95	507	0.152	42.24 ^{+0.42} _{-1.70}	24.46 ^{+0.52} _{-2.62}	1.92 ^{+0.06} _{-0.06}	0.1 ^{+0.2%} _{-0.1%}	0.81 ^{+1.56} _{-0.61}	0.22	CT

bin-wise background estimation), this method remains consistent with Poisson statistics.

3. **Bayesian parameter estimation.** The presented Bayesian framework allows the estimation of parameters where full probability distributions for each parameter are a natural outcome. Constraints such as unphysical regions in parameter space, knowledge from local samples, and information from other studies can be incorporated. For instance, with the Γ -prior we include the assumption that high-redshift AGN behave like local AGN in some regards, and we propagate the uncertainty of redshifts estimates for each source.
4. **Model comparison.** The comparison of models used here overcomes the limitations of current methods. Likelihood-ratio based methods are approximate results in the limit, which can not compare non-nested models. Unlike approximations such as information criteria, the approach is general so that it is unproblematic for model comparison at boundaries.

The implementation overcomes the weaknesses of standard MCMC, namely unknown convergence and multi-modal parameter spaces (see the discussion of methodology in Section 5.2, and also the Appendix 5.3.1 for a specific case). The computational cost is not higher than classical fitting with error estimation or MCMC.

Taking the small step from the MLE-based approach (“C-stat”) to a Bayesian methodology, one might be concerned that the priors influence the result too much. Similarly, one may ask why parametric models are used when no physical model is available. Non-parametric methods would remove the a-priori assumption of a specific model. Often however, physically motivated models *are* available, and the same is true for priors. Similarly to comparing multiple competing models, multiple priors can be tried to test the robustness of the results. In the next Chapter, I will show that with hierarchical Bayes, no priors need to be assumed (Section 6.1.1).

5.4.2 The nature of the obscurer

The best model to describe the sources in this sample is `torus+pexmon+scattering`. This model consists of an intrinsic powerlaw emission, which is absorbed by a toroidal obscurer.

Next to the Compton-scattering provided by the obscurer, an additional Compton reflection contribution (PEXMON) is detected. This component is also clearly required in individual sources, for instance in the spectrum of source 179 presented in Section 5.3.2. In Figure 5.3.3, where the spectrum of `torus+scattering` is shown in the lower left panel, this component is clearly visible in the data through its most prominent feature, the $Fe - K\alpha$ line. As `torus+scattering` already models the Compton scattering and line emission within the well-constrained line-of-sight obscurer, this component must have a different origin. Radiation may be scattered into the line of sight from denser regions of the torus, if a density gradient is assumed. It is worth re-stating that we photo-electrically absorbed the `+pexmon` component, requiring the reflection to occur behind the LOS column density.

Alternatively, the accretion disk may contribute a reflection spectrum as is known from unobscured objects, which is transmitted through the obscurer. In principle, a higher iron abundance is another hypothesis to increase the yield of the line.

Three different absorption models are considered, which differ mainly in the amount of Compton scattering produced outside the line-of-sight due to volume filling (covering). While **wabs** represents a bullet-like blob in the line-of-sight with no Compton scattering, **sphere** and **torus** model a fully and partially open toroidal absorber respectively. The latter two models, computed using Monte Carlo simulations on a constant-density geometry, differ from **wabs** as they consider Compton scattering and K-shell fluorescence. For the full sample, **wabs+scattering** is ruled out by **torus+scattering** (and also **sphere+scattering**), indicating that these differences are important, i.e. that forward-scattered, low-energy radiation and the additional reflection are observed. This is a relevant finding because it demonstrates that high-redshift sources, with relatively few photons compared to well-studied, nearby AGN, are significantly better modelled by a more complex model than commonly used.

The considered obscurer geometries differ in their covering factor, and thus in the strength of the Compton reflection and line emission (see Figure 5.1.2). The **sphere** geometry has the largest reflective area, while **wabs** does not have any Compton scattering; **torus** constitutes an intermediate case. The model comparison shows that the obscurer has neither an extremely large (**wabs+pexmon+scattering**) nor a vanishing opening angle (**sphere+pexmon+scattering**). This result also holds when only the Compton-thick AGN are considered. This conclusion is in line with previous results of Brightman & Ueda (2012). There, for individual sources, the **TORUS** model (with a soft scattering powerlaw) is preferred over e.g. **SPHERE**, via the likelihood of the best fit. In this work, we put these conclusions on a solid statistical footing by applying Bayesian model comparison. It is worth noting that the spherical geometry can also be excluded based on the fact that a soft scattering component is detected, which requires some of the intrinsic radiation to leak out unabsorbed.

However, the bootstrapping results also indicate that there is some diversity in the geometry, as some sources prefer **wabs+pexmon+scattering** over **torus+pexmon+scattering** (in particular $N_H = 10^{23-24} \text{cm}^{-2}$, while Compton-thick objects prefer the latter (see also Figure 5.3.7). This diversity can also be seen in the preference of individual objects in Figure 5.1.3. The **wabs+pexmon+scattering** model is a phenomenological model where the Compton scattering and line emission **PEXMON** and the photo-electric absorption are completely decoupled. This model can emulate the observed spectra almost equally well, especially in the Compton-thin regime. The preference for this model may indicate that the structure just below Compton-thick regime is complicated. In a geometric interpretation, the torus may have a density gradient, and the difference in observed column densities is created by sampling different viewing angles. In this case, the Compton reflection component may originate from denser regions in the torus, which is not modeled in either of the three obscurer models. The geometrical interpretation is illustrated in Figure 5.5.1 (right). Rivers et al. (2013) also detected the reflection component using a model similar to **wabs+pexmon+scattering**. They used archival data from the Rossi X-ray Timing

Explorer satellite which predominantly detects nearby sources ($z = 0.001 - 0.2$). They also conclude that the obscurer is a thick torus rather than a thin disk, but based on the distribution of R values compared to the optical classification (Seyfert 1/2), which is used as a proxy for the viewing angle.

Finally, a component is detected in the soft energies, constructed to be a (unabsorbed) copy of the intrinsic powerlaw emission, but with lower normalisations (shown in Figure 5.3.3, upper right panel). This soft scattering powerlaw can be attributed to Thomson-scattering by ionised material within the opening angle of the torus, scattering the intrinsic spectrum into the line-of-sight (Turner et al. 1997b). This soft component may be confused with other processes such as thermal disk emission or stellar processes. To remedy this, in the sample selection, host galaxy dominated sources have been removed. However, considered just the subsample of $z > 1$, where only the $> 1\text{keV}$ photons enter the observed band, this component is detected with strong evidence (see Table 5.4). The detection of the soft scattering component is in agreement with Brightman & Ueda (2012), who find it to be a ubiquitous feature of obscured AGN. This component could be an outflow or circum-nuclear gas heated by the AGN. However, a preliminary investigation showed that an `apec` model provides a worse description of this component than a powerlaw, indicating that this component may be non-thermal.

Unobscured objects are often well-described by a simple power law. However, they may show Fe lines and Compton reflection originating from reflection off Compton-thick material outside the line of sight, either from the accretion disk or the torus (Awaki et al. 1991). This explains why, e.g. `wabs+pexmon+scattering` provides a good fit here. The torus simulation used may be a good fit as well if `TORUS` had not been constrained to an edge-on view. Brightman & Nandra (see 2011a) fitted the `TORUS` model in a face-on view to unobscured AGN, such that Compton reflection off the torus is modelled self-consistently in the spectrum.

A number of more complicated variations of the best model, `torus+pexmon+scattering`, have been tried, namely (1) linking the opening angle to $\log N_H$ by decreasing it linearly from 60° to 40° from unobscured to Compton-thick sources, (2) making the opening angle a free parameter for each source, and (3) freeing both the opening and viewing angle. These models yield comparable evidence to `torus+pexmon+scattering` and thus are not justified by the available data in the CDFS. Another possible direction would be to test relativistically blurred reflection spectra, which have been used in local AGN to constrain the BH spin. However, this goes beyond the scope of this work.

As a next step, a model that incorporates a density gradient in the torus geometry would be of interest, to self-consistently consider the line-of-sight obscuration and additional Compton-reflection component (`+pexmon`). The density distribution inside the torus is in principle unknown. A starting assumption would be to assume symmetry in the rotation axis, and that the density gradient increases monotonously towards the plane of rotation. The density gradient may be informed by the distribution of the reflection strengths (R values), e.g. from Rivers et al. 2013, or forced to reproduce the distribution of column densities N_H . Constraining the latter is the focus of the next Chapter. Such a model however assumes the strongest possible form of the unification paradigm, in that all

properties (e.g. N_H) are fully determined by the viewing angle.

5.5 Conclusions

In this chapter, I employed a Bayesian framework for analysing X-ray spectra. This methodology is applied to ~ 350 faint, low-count spectra of AGN in the CDFS to infer model parameters. The approach propagates all uncertainties e.g. the Poisson process of collecting counts, or errors in photometric redshifts determination. The novelty of this work however is to apply Bayesian model comparison.

We consider physically motivated models where various geometries – no obscurer, bullet-like blob in the LOS, toroidal and spherical obscurer – are considered. The best model has (1) an intrinsic power law obscured by (2) a constant-density toroid where photo-electric absorption, Compton scattering and Fe-K fluorescence are considered. We detect the presence of (3) an unabsorbed power law associated with Thomson scattering off ionised clouds. Additional (4) Compton reflection, most noticeable through a stronger Fe-K α line, is also found. We find strong evidence against a completely closed, or entirely open, toroidal obscurer geometry.

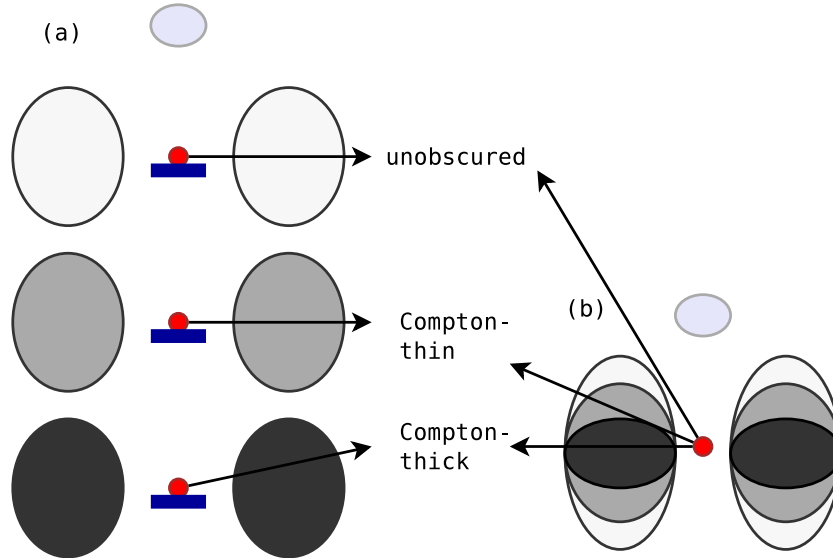


Figure 5.5.1: Cartoon illustrations of a-posteriori possible geometries (see text).

The geometry of the obscurer in the deepest field to date is thus, from the point of view of X-ray spectra, compatible with two simple scenarios illustrated in Figure 5.5.1: (a) Per-source density variations of a constant-density torus, with an accretion disk contributing extra reflection in some sources or (b), following the unification scheme, a torus with a column density gradient where the LOS obscuration depends on the viewing angle and the observed additional reflection originates in denser regions of the torus. In both scenarios, ionised clouds can scatter intrinsic radiation past the torus.

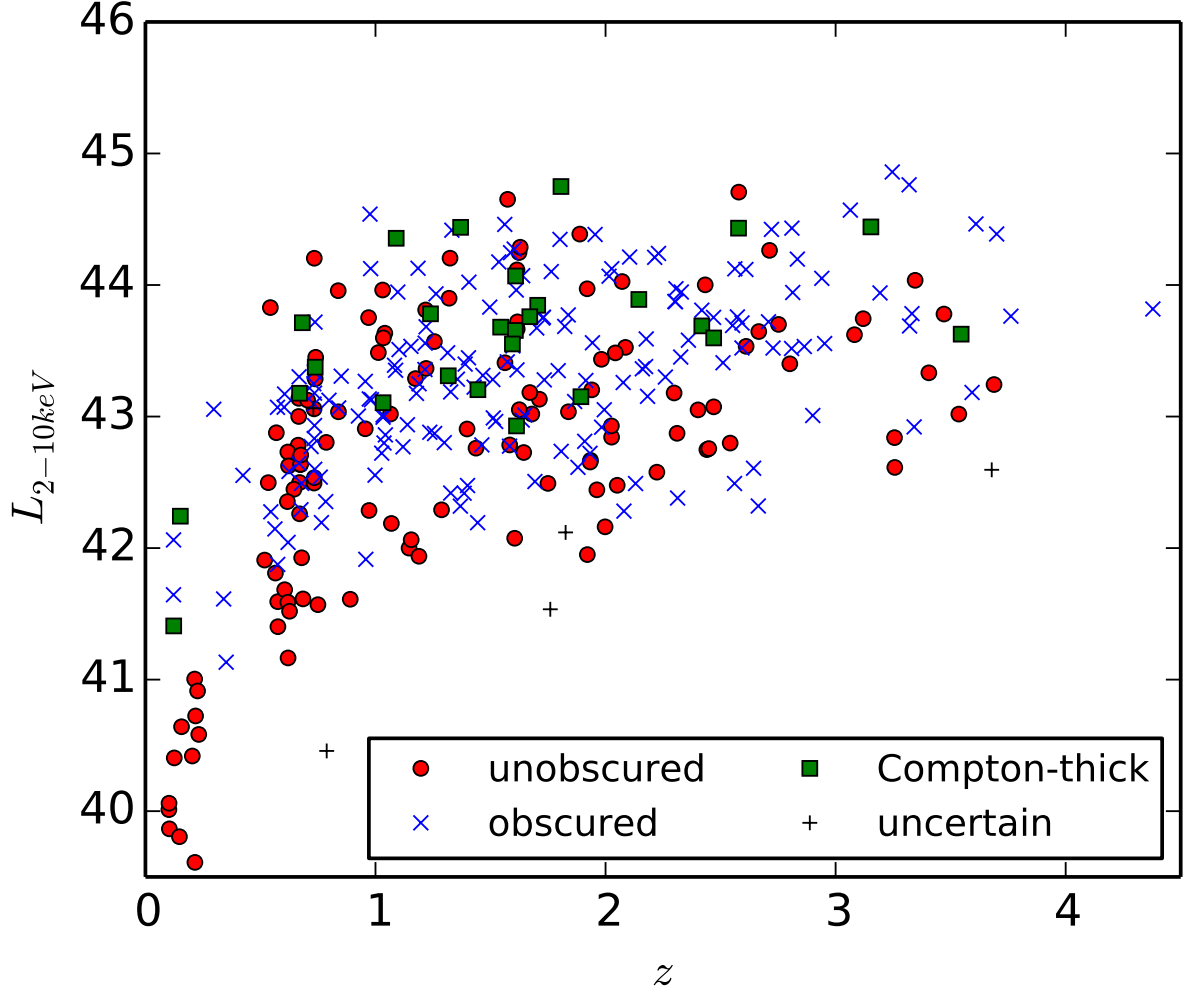


Figure 5.3.6: Luminosity-redshift plot of the sample. The median of the intrinsic luminosity (logarithmic, in erg/s) and redshift posterior probabilities have been used from the `torus+pexmon+scattering` model. Sources are classified as Compton-thick ($N_H > 10^{24} \text{cm}^{-2}$), obscured ($10^{22} \text{cm}^{-2} < N_H < 10^{24} \text{cm}^{-2}$) or unobscured ($N_H < 10^{22} \text{cm}^{-2}$) when the majority of the probability posterior of N_H lies in the respective range. Because of their heavy absorption, the detection of Compton-Thick AGN is biased towards higher luminosities, compared to Compton-thin AGN.

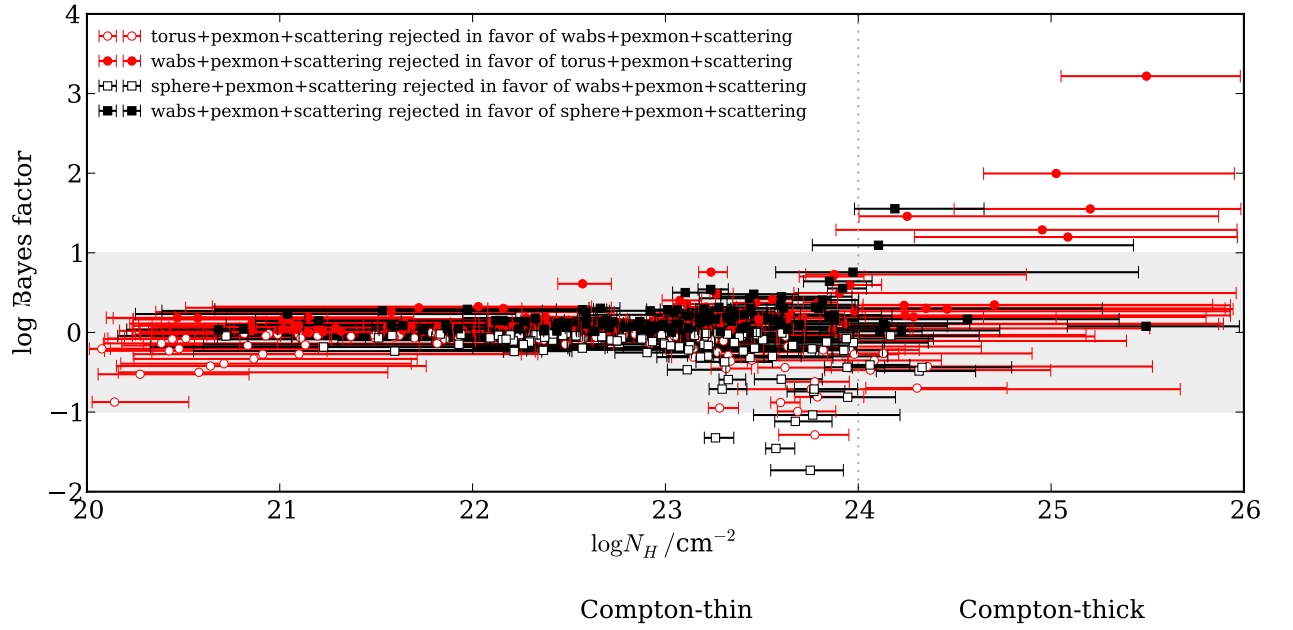


Figure 5.3.7: Evidence contribution from each source with secure spectroscopic redshift. The vertical axis shows the Bayes factor between **torus+pexmon+scattering** and **wabs+pexmon+scattering** (red circles), where strong preference for the torus is above $\log_{10} = 1$. The same is shown for **sphere+pexmon+scattering** and **wabs+pexmon+scattering** (black squares). In both model comparisons, there are obscured objects showing significant preference for each model.

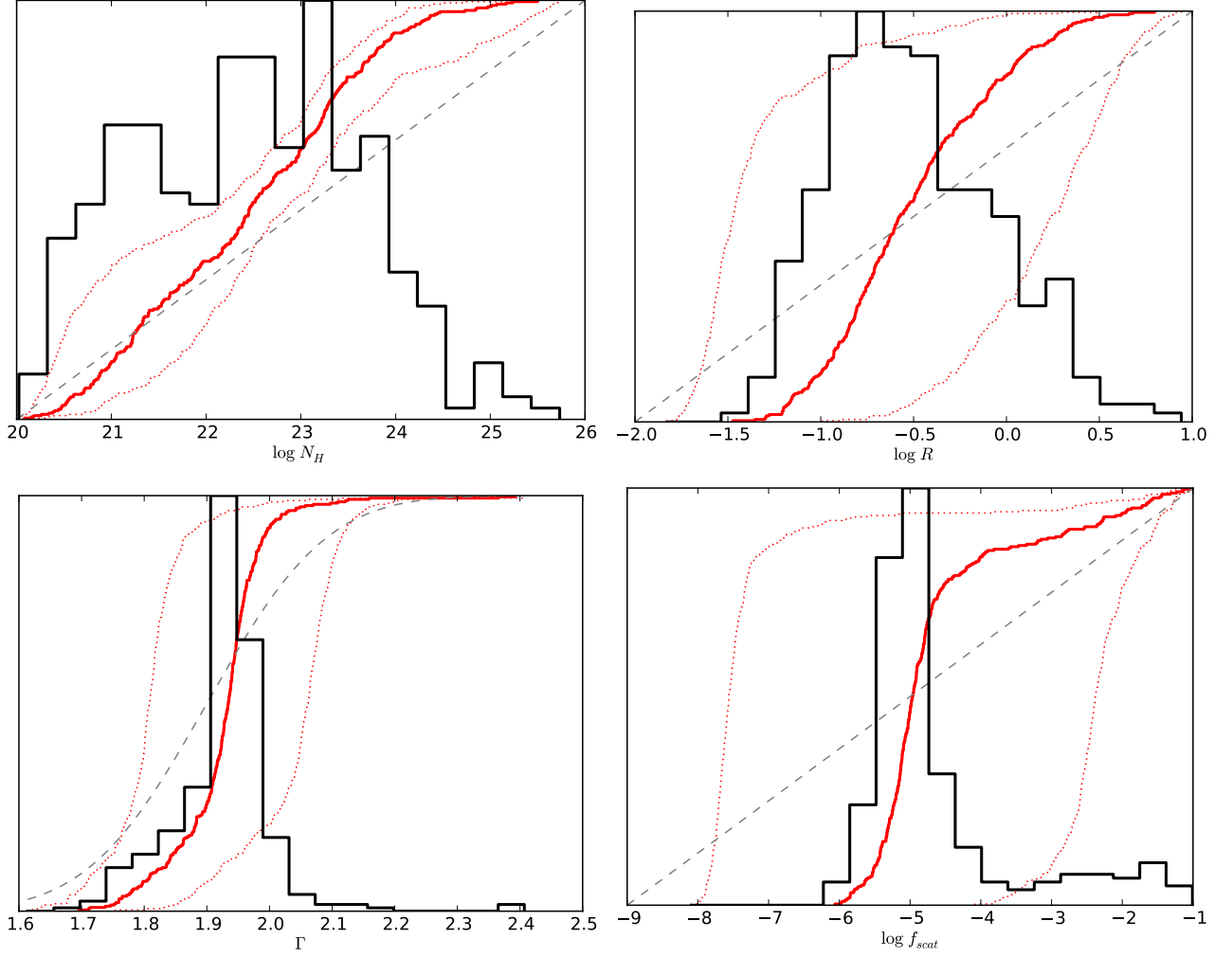


Figure 5.3.8: Histograms of the best parameter values derived using the `torus+pexmon+scattering` model. The median of the marginal posterior distribution for each object is histogrammed in black. The thick red line shows the same as a cumulative distribution. To illustrate the uncertainty in the parameters, the dotted red lines show the cumulative distribution of the 10% and 90% quantiles instead of the median. The dashed gray line shows the assumed prior.

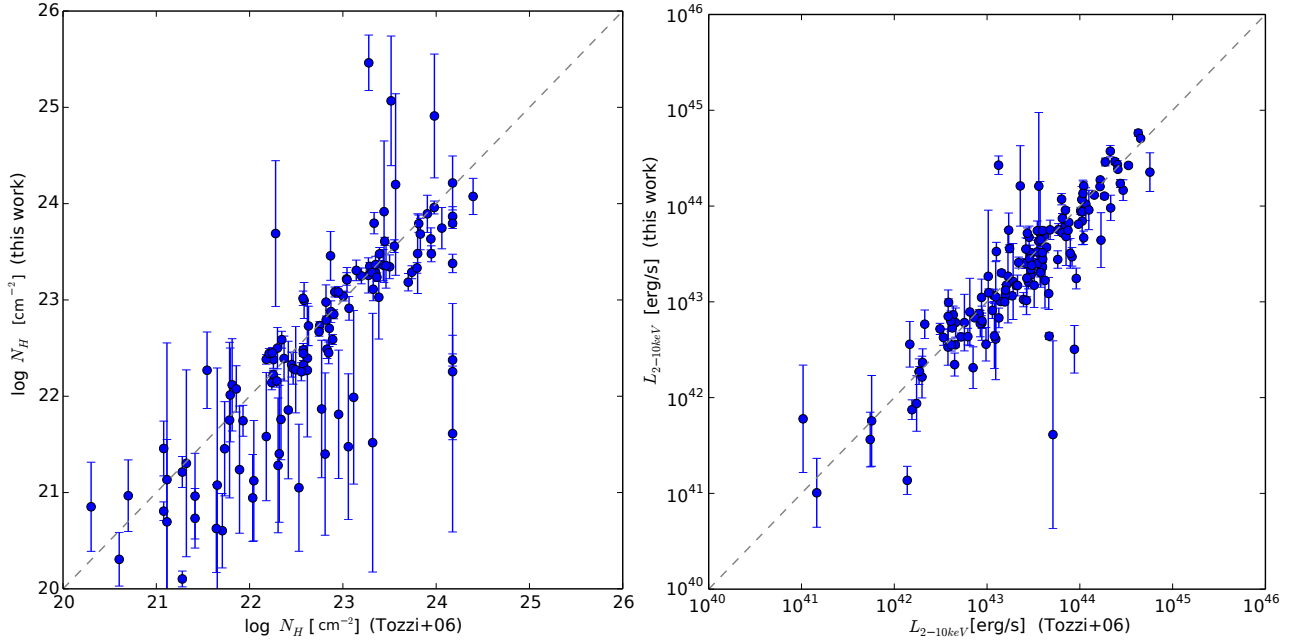


Figure 5.3.9: Comparison of the results from this work with the previous analysis of Tozzi et al. (2006). The *left panel* shows the column density (N_H) while the *right panel* compares the intrinsic luminosity in the 2 – 10 keV rest frame band. Here, only objects are plotted which have the same redshift estimate in Tozzi et al. (2006) and this work. The errorbars show the best fit found in Tozzi et al. (2006) (x-axis) and the median and 1-sigma equivalent quantiles of the posterior in the analysis (y-axis).

6 The relationship between obscurer and accretion luminosity

The first step towards constraining the growth of black holes across cosmic time is to determine the distribution of AGN in accretion luminosity, obscuration and redshift. This is an interesting subject of study because the redshift distribution of AGN shows if the prevalence of AGN has changed over cosmic time, e.g. due to environmental changes such as host galaxy evolution, mergers or disk instabilities (see Section 3.3.4). Furthermore, the X-ray luminosity of AGN is linked to the black hole growth, and thus the distribution of L_X can be translated into growth of black holes (investigated in more detail in the next Chapter). However, to attain these goals it is necessary to constrain the total number of AGN and their total intrinsic luminosity output. This is made difficult by the elusive nature of the most heavily obscured, Compton-thick AGN. The first step is thus to constrain the fraction of obscured and Compton-thick AGN.

Previous work has attempted to constrain the number of Compton-thick AGN using X-ray background synthesis (e.g. Gilli et al. 2007), but while this population is needed to reproduce the shape of the XRB spectrum, Treister et al. (2009) and Akylas et al. (2012) point out that this approach is relatively insensitive to the precise Compton-thick AGN fraction, largely due to the degeneracy with e.g. the assumed level of Compton scattering in spectral models. Multi-wavelength data have also been exploited to identify Compton-thick AGN and constrain their number. These include optical (Risaliti et al. 1999; Cappi et al. 2006; Panessa et al. 2006; Akylas & Georgantopoulos 2009; Gilli et al. 2010; Vignali et al. 2010; Jia et al. 2013; Mignoli et al. 2013; Vignali et al. 2014), infrared (e.g. Fiore et al. 2008, 2009; Alexander et al. 2011; Brightman & Nandra 2011a,b) and hard X-ray (e.g. Sazonov et al. 2008; Burlon et al. 2011; Alexander et al. 2013; Lanzuisi et al. 2014) diagnostics applied to local and non-local AGN samples. These studies estimate Compton-thick fractions relative to the obscured AGN population in the range 30 – 50%, thereby demonstrating that such heavily obscured sources represent a sizeable fraction of the AGN population in the nearby Universe. Following a handful of early discoveries of Compton-thick AGN in deep X-ray surveys (Norman et al. 2002; Tozzi et al. 2006; Comastri et al. 2011), an important recent development has been the identification of significant samples of Compton-thick AGN at moderate to high redshifts (Brightman & Ueda 2012; Georgantopoulos et al. 2013; Buchner et al. 2014; Brightman et al. 2014). This has been enabled by the combination of extremely deep X-ray data, sufficient to constrain the X-ray spectra, along with extensive multi-wavelength coverage of X-ray survey regions, and new techniques able to determine accurate photometric redshifts for X-ray emitting AGN (Salvato et al. 2009, 2011). This then offers the exciting possibility of starting to

constrain the evolution of the obscured AGN populations, including Compton-thick AGN, provided that the selection function can be sufficiently well understood.

Further interest in the obscured AGN population lies in the nature of the obscuration itself, and its possible relationship with other source properties. In the standard unification picture (Antonucci & Miller 1985; Antonucci 1993), all AGN are surrounded by an optically thick toroidal structure relatively close (parsec-scale) to the central engine, which can obscure the line of sight depending on the viewing angle. Alternatively, or additionally, obscuration can occur at galactic scales (Maiolino & Rieke 1995). Observations show that AGN host galaxies are massive and lie, at least in an average sense, on the main sequence of star-formation (Santini et al. 2012). At moderate redshifts, $z \approx 1 - 2$, such galaxies are known to be gas-rich, with gas contents 3 to 10 times larger than local samples (Tacconi et al. 2013). It is therefore possible that obscuration in moderate redshift AGN is associated with the same gas fuelling both star formation and the accretion process itself. In some scenarios (Hopkins et al. 2006a, 2012), obscured AGN represent a distinct phase in the co-evolution of the galaxy and its central black hole, with energy output from accretion sweeping up gas from the surroundings with potentially profound effects on star formation (Silk & Rees 1998; Fabian 1999; King 2003). Determining accurately the fraction of obscured AGN, including Compton-thick AGN, as a function of other parameters such as luminosity and redshift is critical for both unification and co-evolution models.

Previous work has provided evidence that the obscured fraction depends on luminosity, with obscuration being less common in more powerful sources. In a local AGN sample, Lawrence & Elvis (1982) found a correlation between the count rate (luminosity) and hardness ratios (proxy for obscuration). This work has later been confirmed by luminosity function-type studies in the soft (Hasinger 2008) and hard band (Ueda et al. 2003, 2014), as well as in samples detected with higher energies (Burlon et al. 2011). However, the fraction of obscured sources e.g. at the bright end has deviated strongly between works. On the extremes, Hasinger (2008) finds that the fraction goes to zero at $L_X = 10^{45.5} \text{ erg/s}$, while e.g. Treister et al. (2005) claims that the luminosity-dependence is merely a selection effect, in particular for the soft band. A similar luminosity-dependence has been observed in the optical, where the fraction of type-2 AGN is dependent on the X-ray luminosity (e.g. Bongiorno et al. 2010; Steffen et al. 2003). Using data from the COSMOS survey Merloni et al. (2014) finds that the luminosity-dependence of the type-2 fraction is consistent across all redshift bins up to $z \approx 2$.

As the Unification paradigm explains the unobscured/unobscured by viewing angles inside/outside the opening of the torus the most obvious explanation is that the latter increases with luminosity. The receding torus model (Lawrence 1991) is commonly used to describe the luminosity-dependence. In that model, the radius of the torus increases with luminosity, while its height remains the same. This effectively increases the opening angle, providing more unobscured line-of-sights. This model is not a physically consistent setup, as it does not define which physical effect pushes the torus to larger distances. This could be e.g. photon pressure or ionisation of the inner regions. Nevertheless, it provides an empirical fit to the observed behaviour. Simpson (2005) modified the formulae to fit newer data (including those from Ueda et al. 2003; Hasinger 2008), showing that the

initially assumed geometry has to be modified. Mayo & Lawrence (2013) present a model for a partially covered AGN population, in which a significant fraction of the population is covered to 99% by an Compton-thick obscurer. This leads to a spectra with a strongly absorbed powerlaw and another powerlaw component due to the leaking of the intrinsic strength with 1% of the intrinsic luminosity. If spectral fitting with an absorbed powerlaw is applied, the latter component may be mistaken for the intrinsic powerlaw and the source will be identified as an unobscured AGN with a luminosity of 1% of the true luminosity. Such a error leads to a luminosity-dependent fraction of obscured AGN, as the covered population is estimated with a different luminosity than the unobscured population. Such an effect can be problematic in previous works that apply spectral fitting with a simple model (e.g. absorbed powerlaw) or hardness ratios. It is worth emphasising that the new spectral analysis method presented in Chapter 5 does not suffer from this problem: the possibility of a Compton-thick obscurer with a soft scattering component is considered at the same time as the source being unobscured. The probability distribution on the luminosity then contains both solutions (see Section 5.3.1 for an example).

One approach to study whether the geometry of the torus is indeed affected is to observed individual objects in detail. In particular, the infrared-to-bolometric luminosity ratio has been used as a proxy of the covering factor (e.g. Lusso et al. 2013; Roseboom et al. 2013). However, such a derivation strongly depends on the assumed model parameters, such as the inclination, number, size and geometrical distribution of torus clouds, as well as opacity, which are highly degenerate.

The luminosity-function work of La Franca et al. (2005) has allowed in their empirical modeling of the obscured fraction an evolution with redshift. The best fit parameters indeed point to an evolution where in the high-redshift universe, obscured AGN were more abundant. This has remained controversial due to selection effects, as simulated by Akylas et al. (2006). Treister & Urry (2006) consider the optical classification of an X-ray detected AGN sample, and observe no evolution in the observed fraction. However, one would expect less type 2 AGN at high redshift, as they become more difficult to detect. Thus, they argue for an evolution of the obscured fraction. Iwasawa et al. (2012) have analysed a 9 – 20keV (rest-frame) selected sample at $z > 1.7$ and found an elevated fraction of obscured AGN as compared to local samples in the same energy range. Newer luminosity function works also allow for such evolution, however the detection of e.g. Ueda et al. (2014) only shows a very weak increase in the intermediate luminosity range. More specifically, they for example find that $50 \pm 5\%$ of Compton-thin AGN are obscured at $z = 0.1 - 1$ for $L_X \approx 10^{43.8} \text{erg/s}$ AGN while at $z = 1 - 3$, the fraction is $58 \pm 3\%$ at $L_X \approx 10^{44.3} \text{erg/s}$. The weak evidence thus requires further investigation.

Typically, these studies rely on the association of X-ray sources to optical/IR counterparts and subsequent redshift determination. Early luminosity-function type studies in the ROSAT era (e.g. Boyle et al. 1993) only sampled the brightest AGN, which yield high association completeness and easy redshift determination. However, the XMM-Newton and Chandra telescopes have dramatically increased the sensitivity and resolution of X-ray observations. In deep fields, this may lead to the detection of faint X-ray sources without optical/IR counterparts even in deep HST imaging. Additionally, the still relatively large

positional errors can lead to multiple possible counterparts, an ambiguity that often can not be resolved. Also, the faint optical magnitudes of the majority of sources does make spectroscopic followup challenging. These points were initially neglected in early works (e.g. Cowie et al. 1996) until methodological improvements in optical identification (e.g. Brusa et al. 2007) and photometric redshifts (e.g. Salvato et al. 2009) had evolved. The inhomogeneous exposure of XMM-Newton and Chandra observations (as opposed to the relatively even ROSAT exposures) also necessitated advances in the detection of sources and quantification of the selection bias (e.g. Georgakakis et al. 2008). A major advance in methodology was the work of Aird et al. (2010), which accounted for missing counterparts, Eddington bias, Poisson uncertainties in the flux and photometric redshifts. They also used Bayesian model comparison to distinguish between luminosity function models. The luminosity function of Aird et al. (2010) is based on the hard band flux, without accounting for source obscuration or other spectral components. This work continues the methodological advances by incorporating systematic uncertainties in photometric redshift estimation (see Section 4.7.1). The spectrum of each source is modelled in detail to obtain the intrinsic luminosity and obscuration. The uncertainties from unknown counterparts, photometric redshift estimation as well as the limited number of counts are propagated (see Chapter 5).

The most notable contribution of this chapter is the determination of the space density in the $L_X - z - N_H$ space without assuming a parametric model of e.g. the luminosity-dependent behaviour or the evolution of the obscured fraction. The methodology is described in the following section. Chapter 7 explores the shape of the luminosity function and the total number density of AGN over cosmic time. This chapter focuses on the N_H distribution and its luminosity-dependence and evolution.

This work was published as ? in *Astrophysics Journal*, Volume 802, 89 pp.

6.1 Methodology

In the previous chapters I compiled a large sample of hard X-ray selected AGN from the CDFS, AEGIS-XD, COSMOS and XMM-XXL surveys (see Chapter 4 and Section 4.8 in particular), for which intrinsic X-ray luminosity L_X in the 2 – 10 keV energy range, redshift and equivalent hydrogen column densities have been derived. These observables can now be used to infer the distribution of AGN of these key quantities in the population. However, the distribution of the AGN sample analysed in Chapter 5 in intrinsic luminosity, redshift and obscuring column density, is not very useful at the moment, for two reasons:

The foremost problem is the selection bias. X-ray emission provides the least biased way of selecting AGN over a wide range of column densities and luminosities, as X-rays penetrate the obscuring material and the 2–10keV window is largely free of any host galaxy contamination. However, AGN with faint intrinsic emission, those at large cosmological distances and those enshrouded in gas and dust are detected with lower efficiency in flux-limited detections such as the one used in this work, leading to severe incompleteness in these regions of the parameter space. When the column of equivalent neutral hydrogen

exceeds unit optical depth corresponding to the Thomson cross section ($N_H \approx 1.5 \times 10^{24} \text{ cm}^{-2}$), AGN become difficult to find even at X-ray wavelengths. In this “Compton-thick” regime, AGN can nonetheless be identified by hard X-ray emission which can emerge from the thick covering and/or via radiation that is reflected and scattered into the line of sight. These processes result in a characteristic shape to the X-ray spectrum, including a flat continuum and intense iron $K\alpha$ emission lines which can be identified using spectral analysis.

Secondly, the distribution of the observed properties is difficult to use if we want to stay true to the uncertainties. Faint objects, which dominate the sample, are highly uncertain in their properties, namely the intrinsic luminosity L , obscuration N_H , but also their redshift z . For instance, Compton-thick AGN can have large “probability clouds” for their parameters. This prohibits us from assigning objects to bins for visualisation. For direct visualisation, three approaches can be considered: (1) assigning each object to a random luminosity bin based on its probability distribution, and then estimating the density in the bin, (2) assigning each object to each luminosity bin with a probability weight, and then estimating the density in the bin, (3) computing for each luminosity bin the number of objects that have a higher luminosity with e.g. 90% probability. Method (2) has the difficulty that the “number” in each bin is no longer integer – requiring interpolation of the Poisson distribution formula. The methods (1) and (2) assume a frequency interpretation of the uncertainty probability distributions – which is not reasonable as every object is different and the sample size is small. Method (3) may be useful for checking whether data and model agree, but does not yield a intuitive visualisation. None of these methods are satisfying, so here I develop a new approach, which visualises the data and estimates the luminosity function at the same time.

6.1.1 Luminosity function analysis

To obtain the distribution of the underlying population in luminosity, redshift and column density, inference has to take into account that the sample under study is a biased, random draw. Such inference methods have been known for a long time (Marshall et al. 1983). In the following, I review the statistical footing of analysing population demographics. For this work, the usual Poisson likelihood is used:

$$\mathcal{L} = \frac{\prod_k \int \frac{\phi(\mathcal{C})}{d\mathcal{C}} \cdot p(d_k|\mathcal{C}) \cdot \frac{dV}{dz} d\mathcal{C}}{\exp \left\{ \int \frac{\phi(\mathcal{C})}{d\mathcal{C}} \cdot A(\mathcal{C}) \cdot \frac{dV}{dz} d\mathcal{C} \right\}} \quad (6.1.1)$$

where $\mathcal{C} = \{\log L, z, \log N_H\}$, $p(d_k|\mathcal{C})$ represents the results of the spectral analysis of data d_k from the detected object k , which is weighted by the luminosity function ϕ . Usually, a functional form is assumed for ϕ , such as a broken powerlaw, whose parameters are sought (see Figure 6.1.1). The integral in the denominator computes the expected number of sources by convolving the luminosity function with the area curve A (shown in Figure

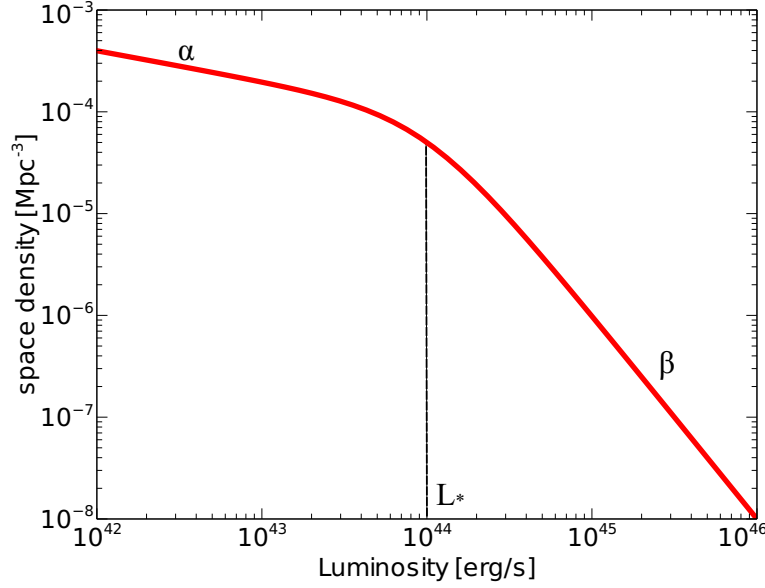


Figure 6.1.1: Broken powerlaw luminosity function model. In this context, a broken powerlaw is defined as $\phi(L) = \frac{N}{\left(\frac{L}{L_*}\right)^\alpha + \left(\frac{L}{L_*}\right)^\beta}$. The parameters are the normalisation N , the break luminosity L_* and the power law slopes above and below the break (α and β).

4.2.2 as a dotted magenta line). The remainder of this section gives an extensive derivation and caveats for this likelihood.

I now go through a derivation of the statistical footing of analysing population demographics by reviewing and combining the works of Loredó (2004) and Kelly et al. (2008). Their main difference is whether the Binomial distribution or its approximation, the Poisson distribution, is used.

The intent is to estimate the number density per comoving volume (Mpc^3) of AGN, as a function of various properties, specifically X-ray luminosity, redshift and obscuring column density. This extended luminosity function will describe the evolution of the X-ray population and all sub-populations (e.g. Compton-thick AGN). The difficulty is that each of these properties influences the ability to detect objects with such properties. Lets assume the probability to detect a source with properties \mathcal{C} can be described as $p(\mathcal{D}, \mathcal{C})$.

After analysing each object in detail, the sample under study could be binned into so small bins that only 1 item at most can be in each bin \mathcal{C}_i . Then just applying the Binomial/Poisson distribution, assuming that in k bins a source is detected (and no detections in the other $n - k$ bins), yields:

$$B(k; n, p) = \underbrace{\binom{n}{k} \times \prod_{i=1}^k p(\mathcal{D}, d_i, \mathcal{C}_i)}_{\square} \times \underbrace{\prod_{i=k}^n p(\bar{\mathcal{D}}, d_i, \mathcal{C}_i)}_{*} \quad (6.1.2)$$

$$P(k; n, p) = \underbrace{\frac{1}{k!} \times \prod_{i=1}^k n \times p(\mathcal{D}, d_i, \mathcal{C}_i)}_{\square} \times \underbrace{\exp \left\{ - \sum_{i=1}^n p(\mathcal{D}, d_i, \mathcal{C}_i) \right\}}_{*} \quad (6.1.3)$$

The first equation gives the likelihood based on the Binomial distributions, and here $p(\bar{\mathcal{D}}, d_i, \mathcal{C}_i)$ denotes the probability of not detecting a source with properties \mathcal{C}_i and having obtained the observed data d_i . In the Poisson formula, only $p(\mathcal{D}, d_i, \mathcal{C}_i)$ occurs, which instead denotes the probability of detection.

Notice that the term marked with a star in Equation 6.1.3 remains the same, regardless of the number of detections k in the sample (only dependent on n). The sum in this term therefore has to be always the same, independent of the specific sample d_i . In the Poisson formalism, the exponent of this term has a specific meaning – it refers to the expectation value of the sample (a a-priori assumption). Thus this dependency of the data can be removed, and be treat it as an a-priori probability. Mathematically, I integrate over all possible data d_i .

$$P(k; n, p) = \exp \left\{ - \sum_{i=1}^n p(\mathcal{D}_i, \mathcal{C}_i) \right\} \times \quad (6.1.4)$$

$$\prod_{i=1}^k (n \times p(\mathcal{D}_i, d_i, \mathcal{C}_i)) \quad (6.1.5)$$

Here I used $\sum_{i=1}^n p(\mathcal{D}_i, d_i, \mathcal{C}_i) = \sum_{i=1}^n p(\mathcal{D}_i, \mathcal{C}_i)$. By analogy, for the binomial distribution, I use $\int p(\bar{\mathcal{D}}_i, d_i, \mathcal{C}_i) dd_i = p(\bar{\mathcal{D}}_i, \mathcal{C}_i)$. Mathematically, integrate over all possible data d_i is done, but for all non-detected sources (see Kelly et al. 2008).

$$B(k; n, p) = \binom{n}{k} \times \prod_{i=k}^n p(\bar{\mathcal{D}}_i, \mathcal{C}_i) \times \quad (6.1.6)$$

$$\prod_{i=1}^k p(\mathcal{D}_i, d_i, \mathcal{C}_i) \quad (6.1.7)$$

$$= \binom{n}{k} \times \exp \left\{ \sum_{i=k}^n \ln p(\bar{\mathcal{D}}_i, \mathcal{C}_i) + \right. \quad (6.1.8)$$

$$\left. \sum_{i=1}^k \ln p(\mathcal{D}_i, d_i, \mathcal{C}_i) \right\} \quad (6.1.9)$$

Now I replace the n discrete bins by a continuum. The probability $p(\mathcal{D}_i, d_i, \mathcal{C}_i)$ can then be non-zero over a range of the parameter space. After all, the true parameter \mathcal{C} are unknown. Thus, I replace $\sum_i p(\mathcal{D}_i, d_i, \mathcal{C}_i)$ by $\int_{\theta} p(\mathcal{D}_i, d_i | \mathcal{C}) d\mathcal{C}$.

$$P(k; n, p) = \frac{1}{k!} \times \exp \left\{ - \int p(\mathcal{D}, \mathcal{C}) d\mathcal{C} + \right. \quad (6.1.10)$$

$$\left. \sum_{i=1}^k \ln \int p(\mathcal{D}_i, d_i, \mathcal{C}) d\mathcal{C} \right\} \quad (6.1.11)$$

$$B(k; n, p) = \binom{n}{k} \times \exp \{ \quad (6.1.12)$$

$$(n - k) \cdot \ln \int p(\bar{\mathcal{D}}, \mathcal{C}) d\mathcal{C} + \quad (6.1.13)$$

$$\left. \sum_{i=1}^k \ln \int_{\theta} p(\mathcal{D}_i, d_i, \mathcal{C}) d\mathcal{C} \right\} \quad (6.1.14)$$

Now I expand $p(\mathcal{D}, \mathcal{C})$ and $p(\mathcal{D}_i, d_i, \mathcal{C})$ using conditional probabilities and discuss the meaning of the occurring terms.

$$p(\mathcal{D}, \mathcal{C}) = p(\mathcal{C}) \cdot p(\mathcal{D} | \mathcal{C}) \quad (6.1.15)$$

$$p(\mathcal{D}, d_i, \mathcal{C}) = p(\mathcal{C}) \cdot p(d_i | \mathcal{C}) \cdot p(\mathcal{D} | d_i, \mathcal{C}) \quad (6.1.16)$$

Here, $p(\mathcal{C})$ is the probability of finding an object with characteristics \mathcal{C} ($\{\log L, z, \log N_H\}$ here). This is essentially the luminosity function of interest. The term $p(\mathcal{D} | \mathcal{C})$ in Equation 6.1.15 denotes the sensitivity to such objects.

In Equation 6.1.16, the second term, $p(d_i | \mathcal{C})$, is related to the spectral analysis. It denotes the likelihood that this data was generated from a source with e.g. luminosity

L and column density N_H . Finally, the probability of detecting this object, given the data and characteristics \mathcal{C} , has to be considered. This is almost surely 1, as having data associated with an object implies having detected it. However, there is a subtlety here that has been overlooked so far. Some astronomers use the sensitivity $p(\mathcal{D}|\mathcal{C})$ – specifically the area curve of where the object was detected – here, in place of $p(\mathcal{D}|d_i, \mathcal{C}) = 1$. But Loredó (2004) makes a strong point arguing that the sensitivity function must not be used here. Why is it done in practise anyways?

The X-ray source catalogues used in this thesis is constructed by extracting counts within an area that corresponds to 70% of the PSF. The background for this region is estimated, and the probability for the background to produce the observed number of counts computed (no-source probability). If a threshold is exceeded, the source is considered to be real and its spectrum is extracted. In the latter step a much larger region is used, 90% of the PSF, in order to include as many source counts as possible. At the same time however, a larger number of background counts is also included. While in the small detection area it was not possible for the background to produce the observed counts, for faint sources it may be possible for the background to produce the observed counts in the larger spectrum extraction region. This is because the background contribution grows linearly with area, but the source contribution almost stays the same. The likelihood of the analysis thus does allow having no flux from a faint source, i.e. luminosity zero. This contradiction to the detection probability stems from the fact that the information that the counts are *concentrated* in detection region has not been used. In this case, $p(\mathcal{D}|d_i, \mathcal{C}) \neq 1$.

Specifically, I can write it as $p(\mathcal{D}|d_i, \mathcal{C}) = p(> k|d_i, \mathcal{C})$, the probability of having more than k counts in the detection region out of the extracted spectrum counts d_i , where k denotes the number necessary for detection at the current position. An approximation to this number is $p(> k|\mathcal{C})$, the probability to produce the number of counts required for a detection at the current position. A further approximation is $\int p(> k|\mathcal{C}) dA/A = p(\mathcal{D}|\mathcal{C})$, which is the area-average sensitivity curve. This explains why some works use the area curve in the data term, even though it should not be done: it is an attempt in fixing a loss of information introduced when the detection and data analysis processes differ.

I now put all the information together and arrive at the relevant likelihoods:

$$B(k; n, p) = \binom{n}{k} \times \left(\int p(\bar{\mathcal{D}}|\mathcal{C}) \cdot p(\mathcal{C}) d\mathcal{C} \right)^{n-k} \times \prod_{i=1}^k \int p(\mathcal{C}) \cdot p(d_i|\mathcal{C}) \cdot p(\mathcal{D}|d_i, \mathcal{C}) d\mathcal{C} \quad (6.1.17)$$

$$P(k; n, p) = \frac{1}{k!} \times \exp \left\{ - \int p(\mathcal{C}) \cdot p(\mathcal{D}|\mathcal{C}) d\mathcal{C} \right\} \times \prod_{i=1}^k \int p(\mathcal{C}) \cdot p(d_i|\mathcal{C}) \cdot p(\mathcal{D}|d_i, \mathcal{C}) d\mathcal{C} \quad (6.1.18)$$

Comparing Equation 9 in Loredó (2004) with Equation 6.1.18 above, the same result has been achieved with the notation key $p(d_i|\theta) = l_i$, $p(\mathcal{D}|d_i, \mathcal{C}) = 1$, $\mathcal{C} = m$, $n = \omega\delta m$ and $p(\mathcal{D}|\mathcal{C}) = \eta(m)$. The derivation is similar, however I started from first principles and continued in a sound and complete way for both the Poisson and the Binomial distribution.

Comparing Equation 5 in Kelly et al. (2008) with Equation 6.1.17 above, I have obtained the same result with the notation key $\mathcal{C} = (L_j, z_j)$, $p(\bar{\mathcal{D}}|\mathcal{C}) = p(I_j = 0|L_j, z_j)$, $p(\mathcal{C}) = p(L_j, z_j|\theta)$, $p(\mathcal{D}|d_i, \mathcal{C}) = 1$ as well as $p(d_i|\mathcal{C}) = 1$, since Kelly et al. (2008) neglects any measurements uncertainties in the derivation, assuming that luminosity and redshift can be determined perfectly.

Finally, I summarise the logarithms of the likelihoods, neglecting constants.

$$\begin{aligned} \ln \mathcal{L}_B &= (n - k) \times \ln \int p(\bar{\mathcal{D}}|\mathcal{C}) \cdot p(\mathcal{C}) d\mathcal{C} + \\ &\quad \sum_{i=1}^k \ln \int p(\mathcal{C}) \cdot p(d_i|\mathcal{C}) \cdot p(\mathcal{D}|d_i, \mathcal{C}) d\mathcal{C} \end{aligned} \quad (6.1.19)$$

$$\begin{aligned} \ln \mathcal{L}_P &= - \int p(\mathcal{C}) \cdot p(\mathcal{D}|\mathcal{C}) d\mathcal{C} + \\ &\quad \sum_{i=1}^k \ln \int p(\mathcal{C}) \cdot p(d_i|\mathcal{C}) \cdot p(\mathcal{D}|d_i, \mathcal{C}) d\mathcal{C} \end{aligned} \quad (6.1.20)$$

The general-purpose framework presented here is advocated by Kelly et al. (2008). It is a general approach for the inference of population demographics based on specific samples. The importance of this step – rather than analysing trends based on sample statistics – and incorporating selection biases, can not be over-stated.

In this work, I restrict myself to the Poisson likelihood. This allows the space density directly (rather than separating sampled volume and probability). For $p(\mathcal{C})$ I insert the luminosity function model $\phi(\log L, z, \log N_H) \times \frac{dV}{dz}$. For $p(\mathcal{D}|\mathcal{C})$ I insert the area curve (see Figure 4.2.2). For $p(\mathcal{D}|d_i, \mathcal{C})$ I use the area curve of the specific field as an approximation (see above). For $p(d_i|\mathcal{C})$ I should use the likelihood of the spectral analysis. However, I did use intermediate priors in the data analysis (see Chapter 5, specifically Section 5.2.4). I thus extend $p(d_i|\mathcal{C}) = \frac{p(d_i|\mathcal{C}) \times p(\mathcal{C})}{p(\mathcal{C})}$ where the nominator is the posterior distribution computed in the spectral analysis, and the prior has to be divided away again. As flat priors have been adopted in $\log L$, $\log N_H$, and z , which are the units of the integral over $\mathcal{C} = \{\log L, z, \log N_H\}$, this division only contributes to the likelihood as a fixed offset, and is not relevant for further analysis. If different intermediate priors had been used, here they would have to be divided away. In fact, I employed a Gaussian prior on the photon index Γ . However, the luminosity functions ϕ I consider are in fact $\phi(\log L, z, \log N_H) \times f(\Gamma)$ where f is the Gaussian prior used, thus I also do not require a correction here.

Regarding the practical evaluation of the likelihood function, I use posterior samples from the spectral analysis as just described. For the first integral in Equation 6.1.20,

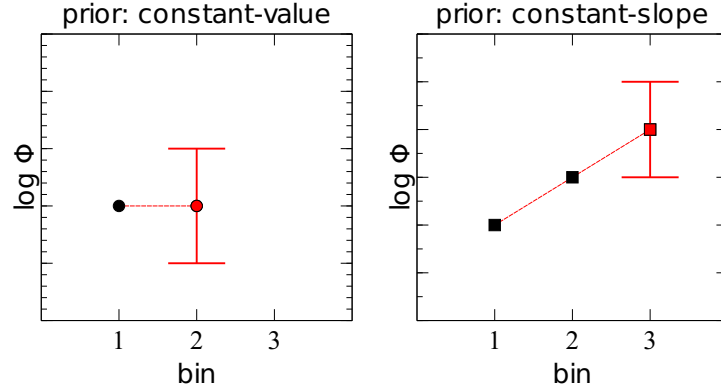


Figure 6.1.2: Illustration of the two smoothness priors used. The *left panel* shows the “constant-value” prior: The extrapolation from well-constrained step-function bins (black) to a neighbouring bin (red) is done by assuming the same value with a fixed uncertainty σ , whose value encodes the assumed correlation strength or smoothness. The *right panel* illustrates the “constant-slope” prior. The value for the neighbouring bin (red) is predicted by continuing the same slope from the black points. As the prior is defined in logarithmic units of the density, this behaviour corresponds to preferring a powerlaw.

which describes the expected number of detected sources, I employ a fixed grid, as these are fast to compute but also ensure a consistent estimate between likelihood evaluations. I use 40 grid points in each of the three dimensions (L , z , N_H). By pre-computation of all weights except for the luminosity function ϕ , the likelihood function reduces to (fast) addition and multiplication operations.

6.1.2 Non-parametric approach

We would like to use the power and safety of a likelihood-based analysis but without the rigidity of a functional form, allowing discovery of the shape of the luminosity function. The method for analysing the luminosity function used here is thus – in a simplistic description – to fit a three-dimensional (L , z , N_H) histogram as the luminosity function model.

Using e.g. $10 \times 10 \times 10 = 1000$ bins already means that the problem is largely under-determined. Thus additional information needs to be provided. Here the reasonable assumption is made that the function does not vary rapidly between neighbouring bins (in particular for the redshift). This smoothness prior is encoded in two approaches, both using the Normal distribution to penalise large deviations. These two priors, “*constant-value*” and “*constant-slope*” are explained below and illustrated in Figure 6.1.2.

- “*constant-value*”: This prior retains the current value unless constraints are imposed by the data. The value of a bin should scatter around its neighbour density value $\log \phi_{i+1} = \text{Normal}(\log \phi_i, \sigma)$ with an allowed correlation width σ for each axis (σ_L , σ_z).

- “*constant-slope*”: This prior keeps power law slopes intact unless constraints are imposed by the data. The log-density slope between bins should scatter around its neighbours slope $\log \phi_{i+2} = \log \phi_{i+1} + \text{Normal}(\log \phi_{i+1} - \log \phi_i, \sigma)$, with the deviation from the slope for each axis (σ_L, σ_z) .

For the N_H dimension, the constant-value prior is always used, as a power law dependence does not seem appropriate here (e.g. Risaliti et al. 1999 finds a approximately flat distribution). With this simple prescription, a smooth field can be recovered by fitting a model whose shape is driven by the data.

Whenever the data constrain the result well, the results from either prior prescription (constant-value and constant-slope) will be the same. Where the constraints are poor, the results will differ depending on the adopted prior. Thus the difference in the reconstructions should be taken as a indication of whether the data or the priors dominate, and in the latter case the difference between the two is an indication of the uncertainty in the determination.

It should be stressed that the choice of binning and in particular the correlation strengths $\sigma_L, \sigma_z, \sigma_{N_H}$ can influence the result. Motivated by the number of data points in each bin, I chose the pixelation as 11 bins of logarithmically spaced luminosity 42 – 46 (units erg/s), redshift bin edges 0.001, 0.1, 0.3, 0.5, 0.75, 1, 1.25, 1.5, 1.8, 2.1, 2.7, 3.2, 4, 7 and $\log N_H$ bin edges 20, 21, 22, 23, 23.5, 24, 26. The correlation strengths σ is defined for neighbouring bins, and their choice is important. Notably, if σ is too small, the model will be flattened out, as the prior dominates, while if σ is too large, no smoothness assumption is used, and uncertainties will be large. Ideally, we would like to recover σ from the data themselves. Unfortunately, the tests show that this computation is not numerically stable. However, above some small value of σ , the results are consistent regardless of the choice of σ . I thus just choose reasonable values for σ , namely $\sigma_L = 0.5$, $\sigma_z = 0.5$ and $\sigma_{N_H} = 0.75$. This encodes, roughly speaking, that neighbouring bins have the same order of magnitude in space density. These values have been chosen after a few initial tests, but were not tuned to give optimal results. Rather, they are one possible, and reasonable, a priori choice.

One aspect of particular interest are the uncertainties of this smooth field reconstruction method. Bins with data will be tightly constrained, while bins without information will have increasing uncertainty with distance from the data. Making uncertainty estimates with so many parameters is not trivial, especially as the parameters are correlated by definition. I use a Hamiltonian Markov Chain Monte Carlo code named “Stan” (Stan Development Team 2014). Stan uses the sophisticated No-U-Turn Sampler (NUTS, Hoffman & Gelman 2011) technique to ensure rapid mixing of the Markov chain by avoiding cyclic explorations. The Stan model used is shown in Algorithm 6.1.

6.2 Results

6.2.1 Obscured and Compton-thick fractions

The total number of AGN, and their distribution in luminosity is the focus of Chapter and presented in Section 7.2.2. The focus of this chapter is the obscuration. Firstly, the

Algorithm 6.1 Stan code for estimating the field (continued in Algorithm 6.2). The constant-slope prior implementation is shown. The integrals to be computed have been prepared so that the density only needs to be evaluated at the respective field bin value (indices) and multiplied by the weights. The evaluation points are chosen using a simple grid, or via the posterior samples of each object.

```

1 data {
2   int indices[27344,3];
3   real weights[27344];
4   vector[11-1] widths0;
5   vector[14-1] widths1;
6   vector[7-1] widths2;
7   int lengths[2046];
8   int chain_indices[780,3];
9   real chain_weights[780];
10 }
11 parameters {
12   real<lower=-40,upper=0> y[11,14,7];
13 }
14 model {
15   real sigmaL;
16   real sigmaZ;
17   real sigmaNH;
18   real dataterms[2046];
19   real detectionterms[780];
20   real loglike;
21
22   sigmaL <- 0.5;
23   sigmaZ <- 0.5;
24   sigmaNH <- 0.75;
25
26   /* L smoothness prior: 2nd derivative is small */
27   for (i in 3:11) {
28     for (j in 1:14) {
29       for (k in 1:7) {
30         y[i,j,k] ~ normal((y[i-1,j,k] - y[i-2,j,k]) * widths0[i-1] / widths0[i-2] + y[i-1,
31           j,k], sigmaL);
32       } } }
33   /* z smoothness prior: 2nd derivative is small */
34   for (i in 1:11) {
35     for (j in 3:14) {
36       for (k in 1:7) {
37         y[i,j,k] ~ normal((y[i,j-1,k] - y[i,j-2,k]) * widths1[j-1] / widths1[j-2] + y[i,j
38           -1,k], sigmaZ);
39       } } }
40   /* NH smoothness prior: 1st derivative is small */
41   for (i in 1:11) {
42     for (j in 1:14) {
43       for (k in 2:7) {
44         y[i,j,k] ~ normal(y[i,j,k-1], sigmaNH);
45       } } }

```

Algorithm 6.2 Continuation of Algorithm 6.1.

```

1  { /* individual objects */
2  int m;
3  m <- 0;
4  for (i in 1:2046) {
5      int a; int b; int c;
6      real v[lengths[i]];
7
8      for (j in 1:lengths[i]) {
9          int l;
10         l <- m + j;
11         /* look up interpolation point and compute weights */
12         a <- indices[l,1] + 1;
13         b <- indices[l,2] + 1;
14         c <- indices[l,3] + 1;
15         v[j] <- exp(y[a,b,c]) * weights[l];
16     }
17     m <- m + lengths[i];
18     /* add value to likelihood */
19     dataterms[i] <- log(sum(v) / lengths[i]);
20 } }
21 /* detection integral */
22 for (k in 1:780) {
23     int a; int b; int c;
24     real v;
25     /* look up interpolation point and compute weights */
26     a <- chain_indices[k,1] + 1;
27     b <- chain_indices[k,2] + 1;
28     c <- chain_indices[k,3] + 1;
29     v <- chain_weights[k] * exp(y[a,b,c]);
30     detectionterms[k] <- v;
31 }
32 loglike <- sum(dataterms) + sum(detectionterms);
33 increment_log_prob(loglike);
34 }
```

CHAPTER 6. THE RELATIONSHIP BETWEEN OBSCURER AND ACCRETION LUMINOSITY

Table 6.1: Key statistics on the fraction of obscured and Compton-thick AGN.

	Obscured fraction ($> 10^{22}\text{cm}^{-2}$)	Compton-thick fraction ($> 10^{24}\text{cm}^{-2}$)
Cosmic time average ^a	$75^{+4}_{-4}\%$	$39^{+7}_{-6}\%$
At $z = 0$ ^{a,b}	$77^{+4}_{-5}\%$	$38^{+8}_{-7}\%$
Maximum ^{a,c}	$83^{+3}_{-3}\%$	$46^{+6}_{-5}\%$
z_{Maximum} ^c	> 2.25	(unconstrained)

^a These fractions relative to the total space density of the population are estimated by integrating the X-ray luminosity function over cosmic time and within X-ray luminosity range $L(2 - 10\text{keV}) = 10^{43.2} - 10^{46}\text{erg/s}$. The uncertainties are computed using the 5% and 95% quantiles of the posterior distribution, i.e. the true value is bracketed with 90% probability.

^b For this estimate the lowest redshift bin is used.

^c The peak is computed by identifying the maximum value (fraction) and location (redshift) in each posterior realisation, and considering the distribution of each series. This leads only an upper limit for the peak location for obscured AGN. For Compton-thick AGN, no upper or lower limit could be determined.

fraction of obscured AGN is reported by comparing the space density above $N_H > 10^{22}\text{cm}^{-2}$ to the total. The non-parametric approach allows us to explore this fraction in a model independent way by integrating the luminosity function over luminosity and cosmic time. These fractions are computed using only the luminosity range $L = 10^{43.2-46}\text{erg/s}$, i.e. without the lowest three luminosity bins. The choice of the luminosity range for the presentation of the results is to minimise uncertainties associated with the typically looser constraints achieved at the faint-end of the X-ray luminosity function. At virtually all redshifts the AGN space density is better determined at luminosities $L > 10^{43.2}\text{erg/s}$. In Table 6.1, the fraction of obscured objects in the Universe is found to be 75%, with narrow uncertainties. The fraction of Compton-thick AGN ($N_H > 10^{24}\text{cm}^{-2}$) is approximately 35%.

Additionally, it is interesting to test whether these fractions are constant over cosmic time. This is done by noting the peak fraction and its redshift in each posterior realisation of the non-parametric reconstruction, and averaging the results. At $z > 2$, the obscured fraction tends to a higher value of almost 85% (see Table 6.1). This shows that the fraction of obscured AGN varies through cosmic time. In the following section, we investigate this evolution of the obscured fraction further. For the Compton-thick fraction, no peak can be identified as the redshift is unconstrained. This indicates that the Compton-thick fraction is constant over cosmic time at approximately 35%.

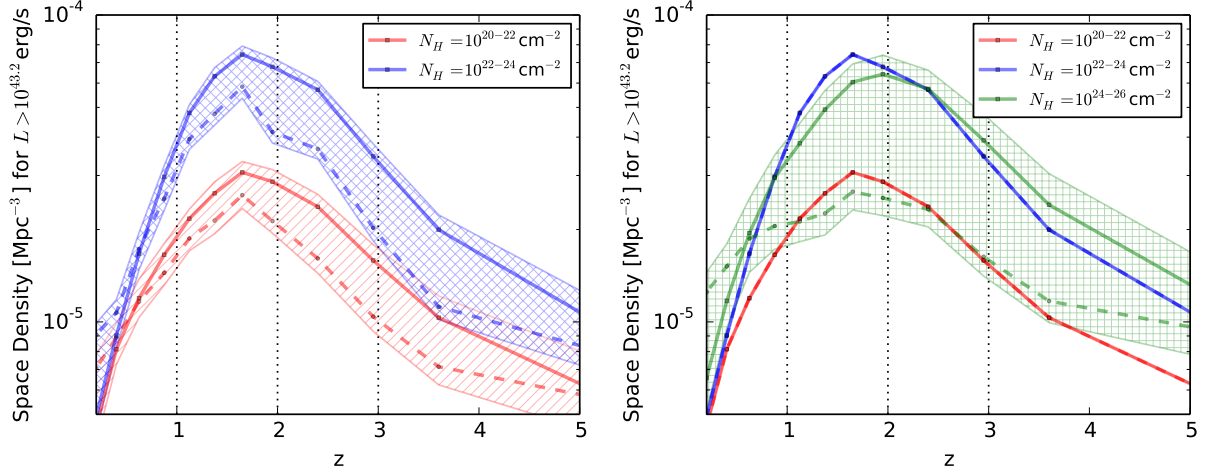


Figure 6.2.1: Redshift evolution of the space density of AGN with $L_X > 10^{43.2}$ erg/s, for various column densities. The step function reconstruction is represented by points at the bin centre, which are connected by lines (dashed for the constant-value prior, solid lines for the constant-slope prior). In the *left panel*, distinct evolutions for the Compton-thin obscured (blue shaded region, top) and unobscured (red shaded region, bottom) can be observed. The *right panel* plots the evolution of Compton-thick AGN as a green shaded region. To facilitate the comparison the evolution of the unobscured and obscured Compton-thin AGN reconstruction is also plotted in the case of the constant-slope prior (solid lines). All AGN sub-populations split by the level of obscuration experience similar space density evolution, which can be described by a rise from $z = 0.5$ to $z = 1.25$, a broad plateau at $z = 1.25 - 2.1$ and a decline at higher redshift. There is also evidence that moderately obscured, Compton-thin AGN ($N_H = 10^{22} - 10^{24} \text{ cm}^{-2}$) are evolving faster in the redshift interval $0.5 - 4$ in the sense that they reach peak space densities higher than the other AGN sub-populations. The space density of Compton-thick AGN has the highest uncertainty, due to poor statistics in the low-luminosity range ($L < 10^{44}$ erg/s). Nevertheless, there is tentative evidence that the evolution of Compton-thick AGN is weaker than that of the Compton-thin obscured AGN (blue), and in fact closer to the evolution of the unobscured AGN (bottom red solid line).

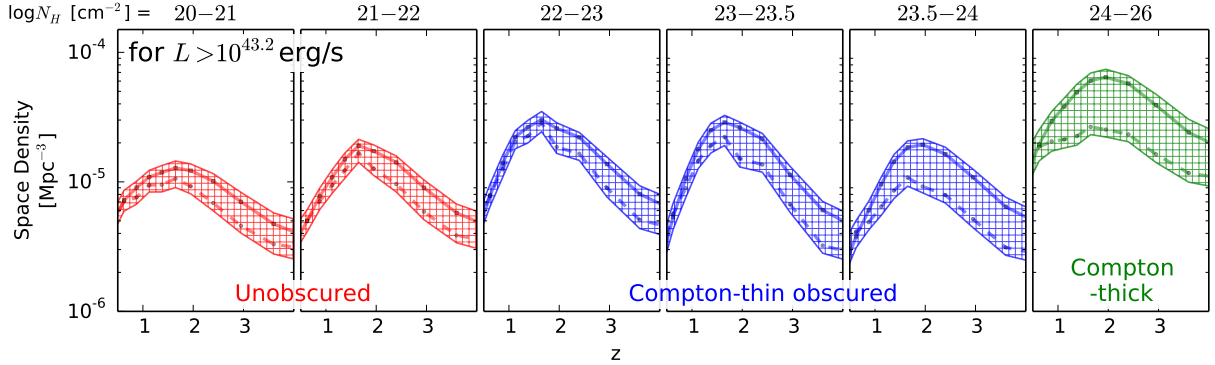


Figure 6.2.2: Redshift evolution of space density of AGN split by the level of obscuration. Different panels correspond to different hydrogen column density interval as indicated at the top. The plot highlights that the redshift evolution of the AGN space density is strongest for the column density bins $N_H = 10^{22} - 10^{23} \text{ cm}^{-2}$ and $N_H = 10^{23} - 10^{23.5} \text{ cm}^{-2}$.

6.2.2 Obscuration-dependent evolution

To explore the evolution of the obscured AGN fraction further, Figure 6.2.1 plots the evolution of the space density of unobscured ($N_H = 10^{20-22} \text{ cm}^{-2}$), moderately obscured ($N_H = 10^{22-24} \text{ cm}^{-2}$) and Compton-thick ($N_H = 10^{24-26} \text{ cm}^{-2}$) AGN with luminosities $L > 10^{43.2} \text{ erg/s}$.

A few words on the visualisation are warranted. All relevant figures plot the posterior distributions of the three-dimensional step function model from various axis views (L , z and N_H). The median result is always shown for the constant-value prior as a *dashed line*, and the median result for the constant-slope prior as a *solid line*. A feature of the non-parametric methodology is that uncertainties are realistic and reflect regions of parameter space where data are sparse. Whenever the data constrain the result well, the results from either prior prescription (constant-value or constant-slope) will be the same. Only when constraints from the data are poor, the results differ. The difference in the reconstructions should thus be taken as an indication of whether the data or the priors dominate the result. The 10%-90% quantile is plotted as *hatched regions* as a measure of the uncertainty, by taking together the posterior samples from both priors.

In the left panel of Figure 6.2.1, moderately obscured and unobscured AGN are found to follow similar evolutionary patterns, namely an increase from $z = 0$ to $z = 1.2$ where their space density peaks and a decline at higher redshifts. However, there are also differences. Moderately obscured AGN evolve much faster from $z \sim 0.5$ to $z \sim 1.5$.

The right panel of Figure 6.2.1 shows the space density of Compton-thick AGN. The evolution of the Compton-thick AGN population has larger uncertainties than unobscured and moderately obscured Compton-thin AGN. Nevertheless the space density shows a broad plateau at $z \approx 1 - 3$, followed by a decline to both lower and higher redshifts. The evolution of Compton-thick AGN also appears weaker than that of moderately obscured

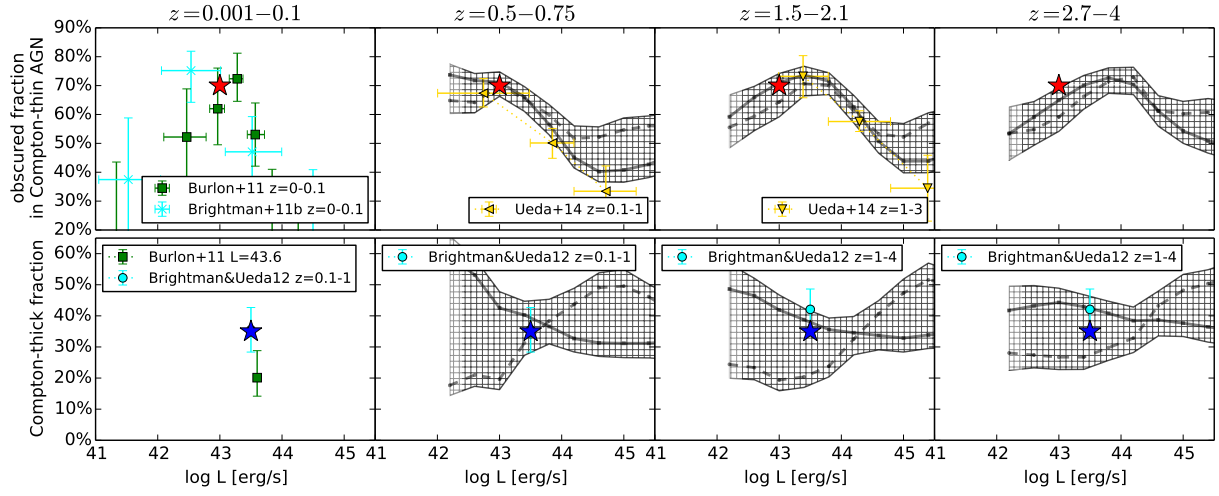


Figure 6.2.3: Luminosity-dependence of the obscurer. The *top row* plots the obscured fraction of Compton-thin AGN (CTNOF, eq. 6.2.1) for various redshift intervals. The shaded grey region are the constraints from the non-parametric method. Additionally, the estimates of Ueda et al. (2014) (yellow points) are shown for comparison. For reference, the red star symbol is placed at 70% and $L = 10^{43}$ erg/s across the panels. In the *top left panel*, the results from local surveys (Burlon et al. 2011; Brightman & Nandra 2011b) which report a similar shape. The CTNOF shows a distinct peak, which is placed at the red star for local surveys, but appears to move to sequentially higher luminosities at higher redshift. In the *bottom row*, the luminosity-dependence of the Compton-thick fraction (eq. 6.2.2) is plotted. The results (shaded grey) show that the Compton-thick fraction is compatible with being constant at $\sim 35\%$ (blue star symbol for reference at $L = 10^{43.5}$ erg/s). For comparison, previous surveys (Burlon et al. 2011; Brightman & Ueda 2012) are shown (see text).

Compton-thin ones. This behaviour is contrary to the N_H smoothness prior, which prefers neighbouring N_H bins to have the same value. Thus it can be concluded that the data drive the result of different evolutions for different obscurations: A strong obscured Compton-thin evolution compared to a weaker evolution in both Compton-thick and unobscured AGN.

The different evolution of AGN with different levels of obscuration is further demonstrated in Figure 6.2.2, where the AGN sample is split into finer N_H bins. All sub-populations experience the same evolutionary pattern, a rise from redshift $z \approx 0.5$, a peak at $z \approx 1.5$ and a decline at higher redshift. The AGN that undergo the strongest evolution are those with column densities around $N_H = 10^{22-23.5} \text{ cm}^{-2}$. Both unobscured and Compton-thick AGN evolve less strongly.

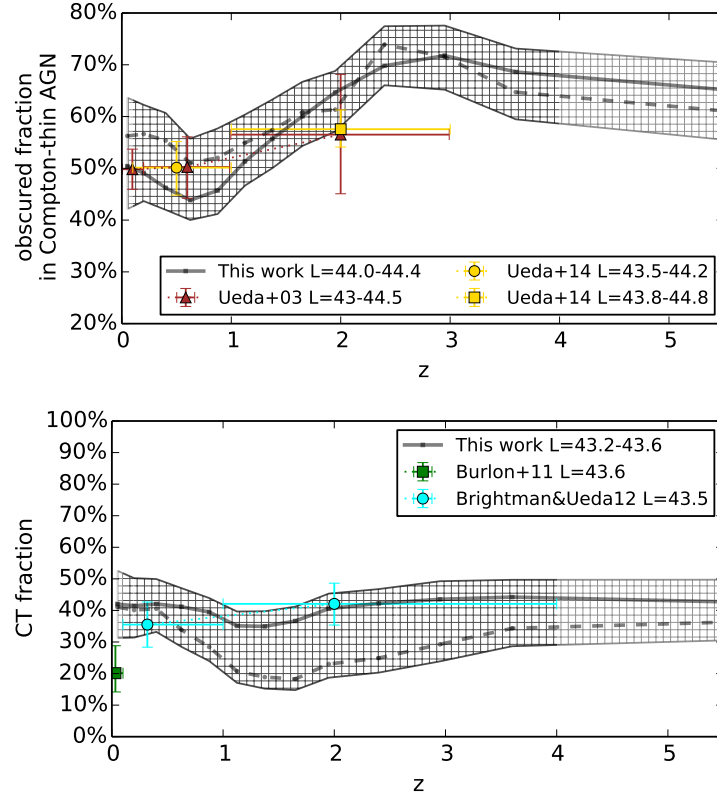


Figure 6.2.4: Evolution of the obscured fraction. *Top panel:* The evolution of the obscured fraction of Compton-thin AGN (CTNOF, eq. 6.2.1) is plotted as a grey shaded region for $L = 10^{44}$ erg/s. At $z \sim 3$, the obscured fraction was higher (75%) than today (50%). For comparison also plotted are the constraints from Ueda et al. (2003) (brown triangles) and Ueda et al. (2014) (yellow points). *Bottom panel:* The evolution of the Compton-thick fraction at $L = 43.5$. The results presented here are compatible with a constant Compton-thick fraction of $\sim 35\%$. I compare to the results to Brightman & Ueda (2012) (cyan circles). Local hard X-ray Swift/BAT observations (green squares, Burlon et al. 2011 for $z < 0.1$) derived a lower Compton-thick fraction of $\sim 20\%$.

6.2.3 Luminosity-dependence and evolution of the obscured fraction

Previous studies suggest that the fraction of obscured AGN is a function of accretion luminosity (e.g. Lawrence 1991; Ueda et al. 2003; Simpson 2005; Akylas et al. 2006; Silverman et al. 2008; Burlon et al. 2011; Ueda et al. 2014). The luminosity-dependence of the obscured fraction is thought to be related to the nature of the obscurer, e.g. reducing it in physical extent. This section investigates this issue using the model independent approach. To illuminate the behaviour of the obscured population I analyse the behaviour of Compton-thin and Compton-thick AGN separately. This section only considers the obscured AGN that are Compton-thin ($N_H < 10^{24}\text{cm}^{-2}$). To this end, a new quantity is defined, namely the obscured fraction of the Compton-thin AGN (Compton-thin obscured fraction, CTNOF)

$$\text{CTNOF} := \frac{\phi[N_H = 10^{22-24}\text{cm}^{-2}]}{\phi[N_H = 10^{20-24}\text{cm}^{-2}]} \quad (6.2.1)$$

In the top row of Figure 6.2.3, the CTNOF appears to be a strong function of luminosity. There is evidence for a peak at a certain luminosity and decline at both brighter and fainter luminosities (the red star symbol provides a constant reference point). Interestingly, the luminosity where the obscured AGN fraction peaks appears to be a function of redshift. With increasing redshift, the drop of the CTNOF at bright luminosities occurs at higher luminosities. These results are in some agreement with the recent analysis in Ueda et al. (2014). One difference however is at the brightest luminosities, where the non-parametric method tends towards a higher value (50%). Similar high obscured fractions were suggested by Iwasawa et al. (2012) and Vito et al. (2014), as opposed to 20% in the local Universe Burlon et al. (2011). Ueda et al. (2014) have better statistics at bright luminosities compared to this work because they included more wide-area and shallow survey fields in the analysis. It is therefore likely that the data used here has poor statistics at the brightest luminosities which makes the preference of the N_H prior towards equipartition apparent. The strength of this sample is rather at the faint-end of the luminosity function for moderate and high redshifts. There, a significant turnover is apparent at low luminosities (right top panel of Figure 6.2.3). This result is independent of the adopted form of the prior, indicating that this behaviour is strongly imposed by the data. A similar behaviour of a peak luminosity and a turnover has been found in local samples (Burlon et al. 2011; Brightman & Nandra 2011b) (top left panel of Figure 6.2.3, corrected assuming a constant 20% Compton-thick fraction in their sample). The position of the peak they find appears to be consistent with the results of this work at low redshift (red star symbol, $L = 10^{43}\text{erg/s}$).

When considering the obscured fraction at $L = 10^{44}\text{erg/s}$, the evolution of the peak is imprinted as a rise with increasing redshift. This is shown in Figure 6.2.4, with the same data points from literature. Observers considering mostly AGN with luminosity $L \geq 10^{44}\text{erg/s}$ will see an increase in the fraction of obscured AGN with redshift as shown here, while observers considering intrinsically faint AGN ($L \leq 10^{43}\text{erg/s}$) would observe

the opposite trend (see Figure 6.2.3).

6.2.4 Evolution of Compton-thick AGN

Compton-thick objects have been hypothesised to play a major role in the accretion phase of AGN. The identification of such sources in current X-ray surveys by XMM-Newton and Chandra is challenging and therefore their space density as a function of redshift has remained controversial. It is therefore important to place constraints on their space density of these sources in the context of previous studies. Figure 6.2.5 shows the evolution of the space density evolution of Compton-thick sources inferred by the non-parametric methodology. Three luminosity cuts are shown which allow direct comparison with previous studies (Fiore et al. 2008, 2009; Alexander et al. 2011), which select Compton-thick AGN at infrared wavelengths. Figure 6.2.5 shows that the general trend is a decline of the space density of this population with decreasing redshift. Above $z = 2$, a decline appears towards increasing redshifts. Also, these results are in rough agreement with previous estimates (Fiore et al. 2008; Alexander et al. 2011). The results of Fiore et al. (2009) on moderate luminosity ($L \geq 10^{43.5} \text{erg/s}$) Compton-thick AGN in the COSMOS field is an exception. The Compton-thick AGN space density determined in that study is significantly higher than the non-parametric estimate from this work. This may be attributed to contamination of obscured AGN samples selected in the infrared by either dusty star-burst or moderately obscured Seyferts (e.g. Georgakakis et al. 2010; Donley et al. 2010).

To study the contribution of Compton-thick AGN to the total accretion luminosity output it is necessary to characterise the luminosity-dependence of Compton-thick AGN. The bottom panel of Figure 6.2.3 plots the Compton-thick fraction, defined as

$$\text{Compton-thick fraction} := \frac{\phi[N_H = 10^{24-26} \text{cm}^{-2}]}{\phi[N_H = 10^{20-26} \text{cm}^{-2}]}, \quad (6.2.2)$$

as a function of X-ray luminosity. Within the uncertainties there is no evidence for a luminosity dependence. If the constant-slope prior is preferred (solid line), a similar behaviour as in the obscured fraction is allowed. However, the results are also consistent with a constant Compton-thick fraction at $\sim 35\%$. This is the first time the luminosity-dependence of the Compton-thick fraction is constrained.

In Figure 6.2.4, the bottom panel shows the evolution of the Compton-thick fraction at $L = 10^{43.5} \text{erg/s}$. A minor dip appears in the $z \sim 1 - 2$ range. This is due to the strong peak in Compton-thin sources (see above), causing a strong rise in the denominator of the fraction (Equation 6.2.2). The blue data point at $z < 0.1$ was taken from the Swift/BAT analysis of Burlon et al. (2011), where a considerably lower Compton-thick fraction was found, in disagreement to the findings here. This may be because of the small volume of the sample at those redshifts, or due to differences in the analysis and in particular the methods adopted to correct for incompleteness of the AGN samples in the Compton-thick regime. Figure 6.2.4 also compares the results of this work with those of Brightman & Ueda (2012) in the Chandra Deep Field South (cyan). Within the uncertainties there is broad agreement. However, Brightman & Ueda (2012) argued in favour of a redshift evolution

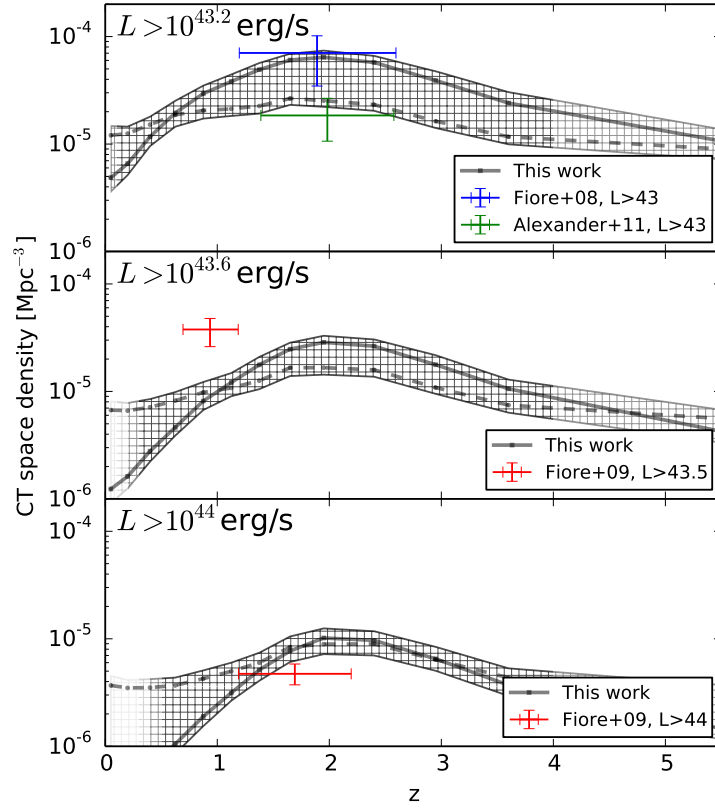


Figure 6.2.5: The evolution of the density of Compton-thick AGN. Different panels correspond to X-ray luminosities brighter than $10^{43.2}$ (*top*), $10^{43.6}$ (*middle*) and 10^{44} (*bottom*) in units of erg/s. Comparing the non-parametric reconstruction (black) to previous results (Fiore et al. 2008, 2009; Alexander et al. 2011), overall good agreement can be seen. The result of Fiore et al. (2009) using the $L > 10^{43.5}$ erg/s cut is an exception. Their estimate lies higher compared to the results of this work, potentially due to starburst galaxy or moderately obscured AGN contaminating their sample (see text).

(increase) of the Compton-thick fraction based on their constraints and those of Swift/BAT (Burlon et al. 2011). Here, there is no evidence for such a trend. The uncertainties in the non-parametric reconstruction are compatible with a constant Compton-thick fraction of $\sim 35\%$.

6.2.5 N_H distribution

This medium luminosity interval, $L \approx 10^{43.5}$ erg/s, where most of the evolution occurs, is also where the obscuration peaks according to Figure 6.2.3. To further illustrate the evolution there, Figure 6.2.6 plots the intrinsic N_H distribution for three redshift intervals in panels. Here, the boxes indicate the upper and lower 90% quantile on the fraction of sources in the respective N_H bin. The dashed line illustrates a possible N_H density distribution that fits these fractions. When comparing the top panel to the middle panel in Figure 6.2.6

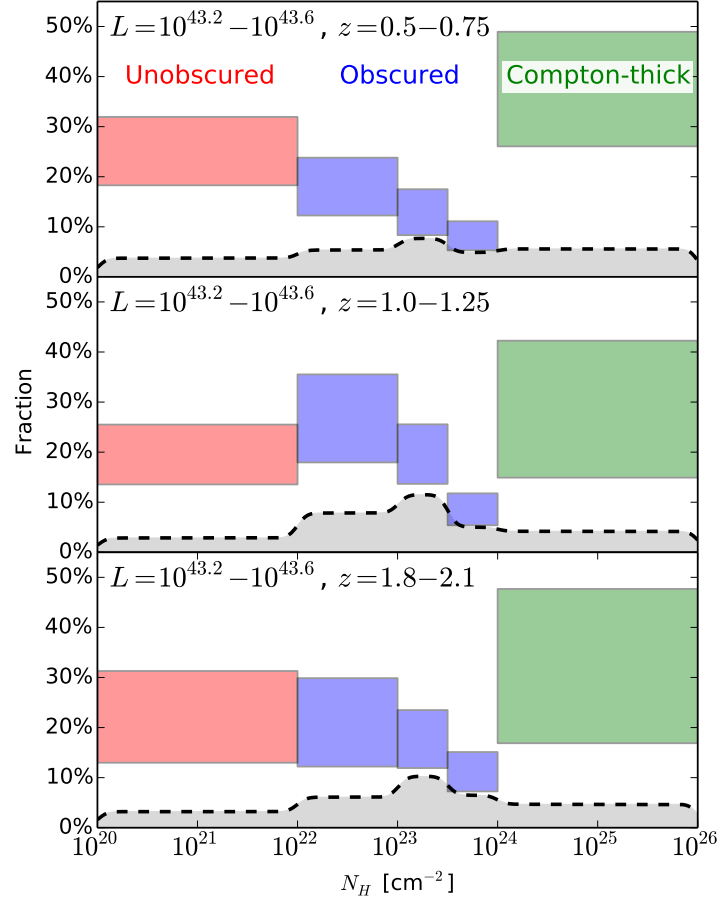


Figure 6.2.6: Column density distributions at various redshifts. Based on the field reconstruction, the fraction of sources in each column density bins (red: “unobscured”, blue: Compton-thin obscured, green: Compton-thick) is computed. The boxes show the credible regions containing 90% of the posterior probability. The dashed line illustrates a possible N_H density distribution that fits these fractions.

in particular, the main effect appears to be that the fractions at $N_H \approx 10^{23} \text{cm}^{-2}$ increase towards higher redshifts.

6.3 Discussion

This work combined deep and wide-area X-ray surveys conducted by Chandra and XMM-Newton to initially constrain the space density of X-ray selected AGN as a function of accretion luminosity, obscuring column density and redshift. The methodology used accounts for the different sources of errors. For example, the analysis includes uncertainties, both random and systematic, associated with photometric redshift measurements as well as the lack of redshift information for sources without optical or infrared counterparts (see Section 4.5 and 4.7). For the determination of the obscuration and intrinsic luminosity of individual sources the X-ray spectral analysis method of Buchner et al. (2014) is used (Chapter 5), which takes into account both the Poisson errors of X-ray spectra and photometric redshift errors. The spectral analysis also uses a physically meaningful, multi-component model that is consistent with recent ideas and observations on dominant AGN emission processes and the structure of matter in the vicinity of active supermassive black holes. Another important feature of the analysis presented in this work is the non-parametric method developed to determine the X-ray luminosity function of AGN. This allows exploration of the space density of AGN as a function of luminosity, redshift, and column density without imposing any model. This frees us from any assumptions on the dependence of the AGN space density to luminosity, redshift, and column density.

Obscured AGN ($N_H > 10^{22} \text{cm}^{-2}$) are found to dominate the population of active supermassive black holes at all redshifts. This is in agreement with previous investigations (e.g. Ueda et al. 2003; Akylas et al. 2006; Ueda et al. 2014; Merloni & Heinz 2013), although the analysis here highlights this point more robustly and quantitatively (with uncertainties). About 75% of the AGN space density, averaged over redshift, corresponds to sources with column densities $N_H > 10^{22} \text{cm}^{-2}$ (see Table 6.1). The bulk of the AGN population is therefore hidden behind large column densities of gas and dust clouds.

6.3.1 The role of Compton-thick AGN

The analysis in this chapter was able to constrain the number fraction of the most heavily obscured, Compton-thick sources to the AGN population, finding it to be $38^{+8}_{-7}\%$ of the total population. Results from previous AGN surveys have been divergent due to difficulty of identifying Compton-thick AGN (e.g. 15 – 20% in Akylas & Georgantopoulos 2009, 5 – 20% in Burlon et al. 2011 for the analyses of Swift/BAT surveys), and relatively low compared to the requirements from the X-ray background (see below). The constraints on the Compton-thick fraction presented here are in good agreement with the estimates of Brightman & Ueda (2012) of $\sim 35 - 40\%$. However in their work they concluded an evolution of the Compton-thick fraction by contrasting their high-redshift data to a local survey, which reported a significantly lower fraction (specifically Burlon et al. 2011). In

this analysis, no evidence for a redshift evolution of the Compton-thick fraction is seen. However, the sample used here lacks large, shallow fields that can probe the local universe, and thus the sample may also benefit from being combined with a local estimate. For instance, the 2 – 10 keV X-ray based work of Risaliti et al. (1999) estimated that 50% of all Seyfert 2 are observed with a Compton-thick line of sight. At face value, this is in agreement with the results found here, without the need for any evolution. However, hard X-ray surveys have reported significantly lower values, e.g. $20^{+9}_{-6}\%$ in Burlon et al. (2011), 9–17% in (Bassani et al. 2006; Malizia et al. 2012; Vasudevan et al. 2013). This shows that estimates for a local Compton-thick fraction have been diverse, making it difficult to make any claim of a evolutionary trend. One possible source of uncertainty is the sensitivity to sources with $N_H = 10^{25-26}\text{cm}^{-2}$. This work has explicitly assumed that their space density is the same as those of sources with $N_H = 10^{24-25}\text{cm}^{-2}$. The sparse sampling (only 3 secure objects in the entire sample) of this heavily buried population prohibits strong inferences.

For discussing the accretion luminosity onto obscured AGN it is noteworthy that the fraction of Compton-thick AGN does not show a luminosity-dependence (Figure 6.2.3). When considering currently accreting AGN, Compton-thick AGN appear to play an invariant role with luminosity and cosmic time. However, the large uncertainties in the luminosity-dependence do not allow firm conclusions. As Figure 6.2.1 shows, the evolution of Compton-thick AGN may follow the unobscured AGN closely in shape, rather than the Compton-thin obscured AGN which evolve strongly. In the former case Compton-thick AGN can be modelled as $35\%/25\% = 1.4$ times more abundant than unobscured AGN over cosmic time and luminosity. In the latter case, they would be approximately equal in abundance and follow the luminosity dependence and its evolution (see Section 6.2.3, discussed below in Section 6.3.3).

The analysis presented places, for the first time, tight constraints on the Compton-thick contribution. Results on the Compton-thick fraction from studies of the cosmic X-ray background have varied between 9% (Treister et al. 2009), 45%-30% (Gilli et al. 2007, luminosity-dependent) and 28%-60% (Ueda et al. 2014, not constrained), while Shi et al. (2013) reported 38%. This scatter is also due to the fact that XRB fitting involves other parameters which are degenerate with the Compton-thick fraction (e.g. additional reflection, see Akylas et al. 2012). It is worth emphasizing that the estimate on the Compton-thick fraction derived here is higher than previous survey studies, but in agreement to those values assumed in X-ray background synthesis analyses.

Having constrained the Compton-thick fraction well, one can speculate on the origin of this obscuration. The fraction of Compton-thick sources is considerably large (35%), necessitating that a large fraction of viewing angles is obscured ($\sim 20^\circ$, see Figure 6.3.1). Compton-thick obscuration is associated with the torus, as such column densities are not typically reached by galactic gas. As a Compton-thick, smooth obscuration would be unstable (Krolik & Begelman 1988), the current working hypothesis is that the obscuration comes in discrete clouds. Under this clumpy torus model obscured views are produced when clouds are encountered in the line of sight. If the Compton-thin obscuration is also associated with the torus, Compton-thick AGN are a simple extension of Compton-thin obscured AGN. In principle it would also be possible then that Compton-thick AGN have

just a larger number of clouds in a clumpy torus. In this case, their cumulative line-of-sight obscuration would provide a Compton-thick view. Alternatively, these objects may have denser clouds overall. Another possibility is that the cloud density is distributed such that both regimes are covered. For further research in this regard, the unbiased column density distributions shown in Figure 6.2.6 can provide observational constraints for clumpy torus models. The alternative scenario is that Compton-thin obscuration is associated with a medium of different extent than Compton-thick obscuration (e.g. through galactic and nuclear gas, Matt 2000). Both kind of models are discussed in Section 6.3.3 when also considering the luminosity dependence.

6.3.2 Obscuration-dependent evolution

Figure 6.2.1 shows that the space density of obscured Compton-thin AGN experience a much stronger rise in the redshift interval $z = 0.5 - 1.2$ compared to unobscured AGN. This demonstrates an observed obscuration-dependent evolution. The non-parametric method used here explicitly contains a smoothness prior, preferring that the space density of AGN as a function of column densities is the same. In contrast, the results show strong differences in the evolution of unobscured and obscured Compton-thin AGN. It is thus safe to conclude that this result was driven by the data.

Figure 6.2.2 indicates that the fastest evolving population is the one with column densities $N_H = 10^{22} - 10^{23.5} \text{cm}^{-2}$. A possible interpretation of these trends is that different levels of obscuration correspond to media with different spatial extents. This is investigated further by focusing on the behaviour of moderately obscured AGN ($N_H = 10^{22} - 10^{24} \text{cm}^{-2}$), discussed in the next section.

6.3.3 Luminosity-dependence and evolution of the obscured fraction

An important result from this analysis is that the obscured Compton-thin AGN fraction (CTNOF, eq. 6.2.1) depends on both luminosity and redshift (see Figure 6.2.3). The luminosity dependence can be described by a peak of the CTNOF at a certain luminosity followed by a decline at both brighter and fainter luminosities. The redshift dependence is manifested by a shift to brighter luminosities of the peak of the obscured AGN fraction with increasing redshift. The shift of this peak causes observers considering mostly AGN with luminosity $L \geq 10^{44} \text{erg/s}$ to see an increase in the fraction of obscured AGN with redshift (as shown in Figure 6.2.4). Previous studies also find that the fraction of obscured AGN decreases with increasing luminosity (e.g. Ueda et al. 2003; Akylas et al. 2006; Ueda et al. 2014). When a luminosity range is considered where the sample has good constraints ($L = 10^{43.2-46} \text{erg/s}$), the obscured fraction rises with redshift so that at $z > 2.25$, $83^{+3}_{-3}\%$ of AGN are obscured ($N_H > 10^{22} \text{cm}^{-2}$), as compared to the local $z = 0$ value $z = 75^{+4}_{-4}\%$.

The results also establishes a decline of the obscured AGN fraction at low luminosities. This trend has only been found in the local Universe (Burlon et al. 2011; Brightman &

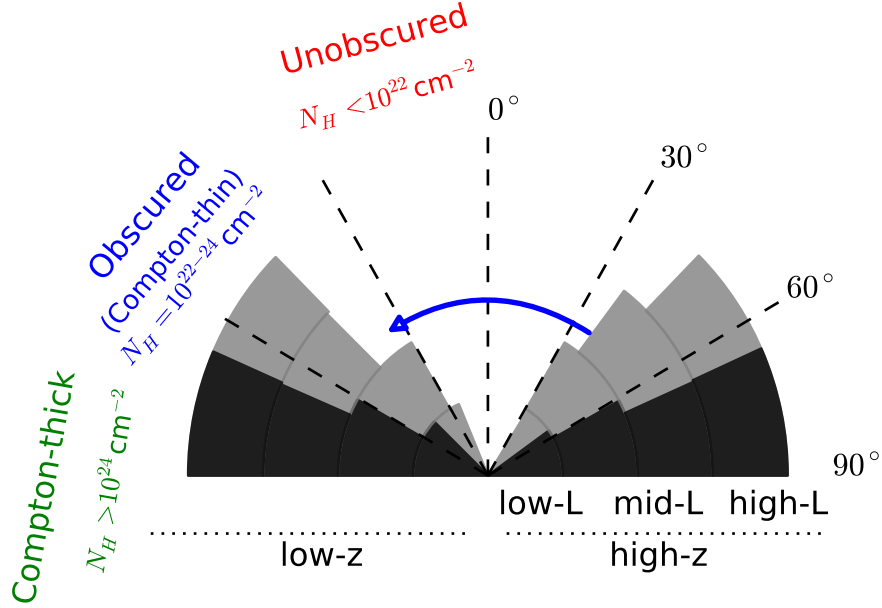


Figure 6.3.1: Illustration of the population-averaged obscuring toroid. The obscured and Compton-thick fraction and their luminosity dependence is shown using a corresponding opening angle above which the sky is obscured to the X-ray emitter (center). The luminosity dependence is depicted for low-redshift (left) and high-redshift (right) AGN using radial shells: as the luminosity increases, more sky is visible. Notice that the difference between a low-redshift and high-redshift source does not affect the lowest and highest luminosities, but only where the transition occurs (blue arrow). The opening angle corresponding to the Compton-thick transition remains roughly constant. For accurate numbers, refer to Figures 6.2.3 and 6.2.4.

Nandra 2011b). Figure 6.2.3 also leads to the conclusion that observers who would consider only intrinsically faint AGN ($L \leq 10^{43} \text{ erg/s}$) would observe a decrease in the fraction of obscured AGN with redshift. In other words, the magnitude of the trend is determined by the sample selection.

Obscurer models that do not invoke any obscuration-dependence are ruled out by the results presented in this work. These include the simple torus where the obscuration distribution is produced purely geometrically. Despite these shortcomings, this picture may serve as a visualisation of the observed trends. The luminosity-dependence and evolution of obscured/Compton-thick AGN is illustrated in Figure 6.3.1. If it is assumed that a toroidal obscurer geometry is shared by the entire population, one can define an opening angle that reproduces the obscured/Compton-thick fraction. The luminosity dependent increase of this opening angle from 30° to 40° is depicted using radial shells for low-redshift (left) and high-redshift (right) AGN. Notice that the difference between a low-redshift and high-redshift source does not affect the lowest and highest luminosities, but only where the transition occurs (blue arrow). Thus, it is not the case that high-redshift sources merely have more obscuration overall. If this were the case, a boost in the obscured fraction would

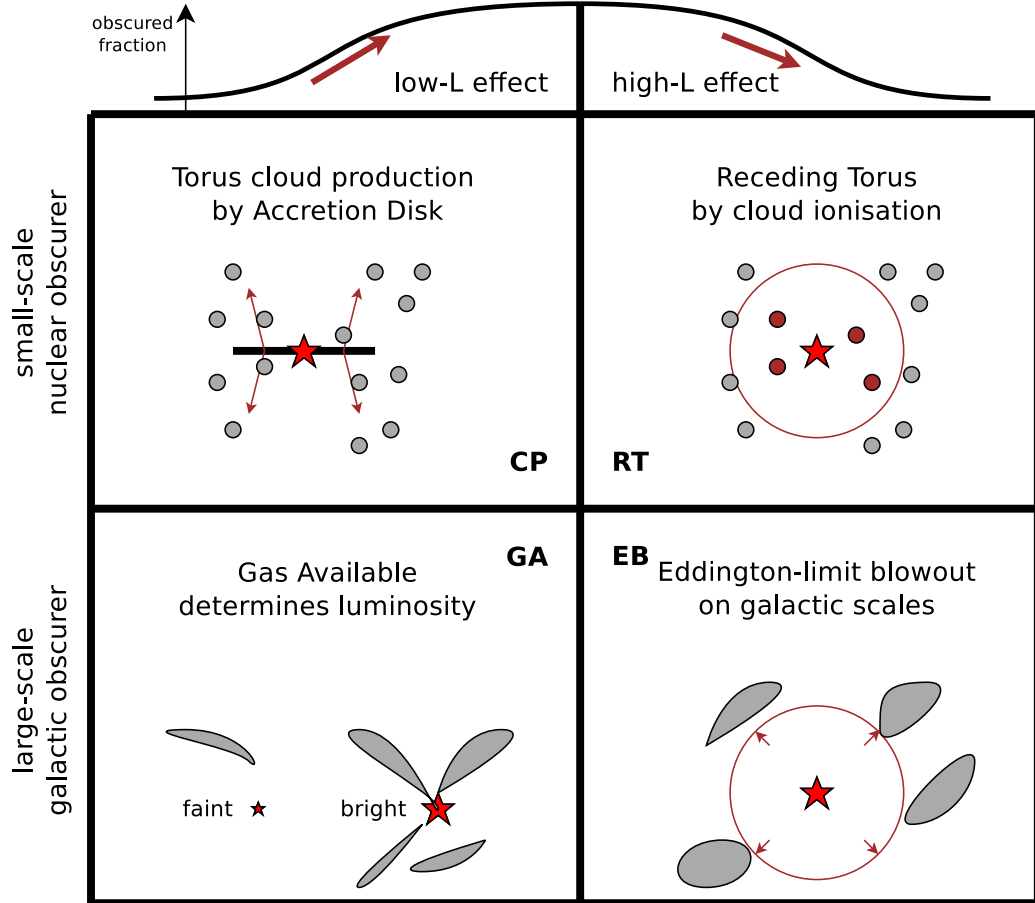


Figure 6.3.2: Four models that address the luminosity-dependence of the obscured fraction. This categorisation distinguishes by the luminosity-dependent effect, the increase of the obscured fraction (left column) and the decrease of the obscured fraction (right column). The top of Figure 6.2.3 is repeated here and simplified. In the four models considered (see text for details), the relevant obscurer is either the torus (top row) or galactic obscuration (bottom row). In the illustration of the models, the red star represents the X-ray source, while grey clumps illustrate the cold obscurer.

be visible over all luminosities. Instead the peak of the obscured fraction has moved to lower luminosities, thus reducing the obscurer in moderately bright AGN (see Figure 6.2.3 for accurate numbers).

Lets now consider a few exemplary models that can qualitatively account for the decrease of the obscured fraction towards low luminosity or the decrease towards high luminosities. An additional uncertainty is whether the relevant obscuration occurs on small or large scales, i.e. in the torus or on galactic scales. In the following, I present for both scenarios a model each for the high-luminosity and the low-luminosity effect. These are illustrated in Figure 6.3.2.

The receding torus (RT) model is often invoked to explain the observed decrease of the obscured AGN fraction at bright luminosities. This scenario postulates a toroidal geometry for the obscurer with opening angle that increases with increasing luminosity. This is attributed to photon pressure pushing away the obscuring material, photo-ionisation of the gas clouds or sublimation of the dust by the photon field of the active black hole (Lawrence 1991; Simpson 2005; Akylas & Georgantopoulos 2008). Such a variation of the scale height of the torus has been claimed via infrared observations (Maiolino & Rieke 1995; Lusso et al. 2013; Toba et al. 2014). In the context of clumpy torus models (Hönig & Beckert 2007; Nenkova et al. 2008), a more intense radiation field leads to the (partial) ionisation of individual dust and gas clouds thereby increasing the effective solid angle of unobscured sight-lines. This model (RT) thus represents a direct, causal connection between the X-ray luminosity and the unobscured line of sight. This scenario is illustrated in the top right of Figure 6.3.2. In the RT model it is, however, difficult to explain the observed evolution of the luminosity dependence. The main evolutionary effect observed is an increase of the turnover luminosity with redshift, causing the onset of the high-luminosity effect at subsequently higher luminosities (see Figure 6.2.3). For a purely nuclear obscurer scenario such as RT, no such evolutionary effect is predicted.

Alternatively, it is also possible to interpret the observed luminosity dependence of the obscured AGN fraction in the context of evolutionary models, in which different levels of obscuration roughly correspond to different stages of the growth of supermassive black holes. Most black hole formation models include an early stage of gas inflows caused by mergers (e.g. Hopkins et al. 2006a; Somerville et al. 2008) or secular processes (e.g. Fanidakis et al. 2012; Bournaud et al. 2007; Ciotti & Ostriker 1997, 2001). During this period super-massive black holes grow fast by accreting material close to the Eddington limit and are also typically obscured by the inflowing material. Eventually however, the energy output of the central engine becomes sufficiently powerful to drive outflows, which can blow away the obscuring clouds of dust and gas. The central engine then shines unobscured for a brief period before its luminosity output declines as a result of the depletion of the available gas reservoirs. It is further proposed (e.g. Hopkins et al. 2005b, 2006c) that luminous AGN brighter than the break of the X-ray luminosity function are dominated by fast accreting black holes close to the peak of their growth phase. Lower luminosity sources include a large fraction of AGN during the decline stage of their activity.

In the above scenario the decrease of the obscured AGN fraction with increasing luminosity may be linked to the blow-out stage of AGN. This scenario (EB) illustrated in the

bottom right panel of Figure 6.3.2. At brighter accretion luminosities it is more likely that AGN-related feedback mechanisms become more efficient, thereby shedding the gas and dust cocoons in the vicinity of supermassive black holes. Lower accretion luminosities, below the break of the X-ray luminosity function, include a large fraction of AGN past the peak of their activity, when the gas reservoirs that feed and obscure the black hole have already been depleted. Therefore, one might expect a larger fraction of unobscured AGN at low accretion luminosities. The evolutionary picture outlined above is therefore consistent with the observed luminosity dependence of the obscured AGN fraction: a maximum at a given luminosity and a decline at both brighter and fainter luminosities.

In the EB model, the observed redshift dependence of the turnover luminosity is also difficult to explain. Unless the physical conditions of the obscuring material are a function of cosmic time, it is not obvious how the physical mechanisms that affect the geometry of the obscurer on small-scales can produce a luminosity-dependent obscured fraction that changes with redshift.

A rescuing argument for the EB scenario is possible. It is now fairly established that the black hole mass of the average, currently accreting AGN increases with redshift (Merloni 2004; Merloni & Heinz 2013). This means that at high redshift, the typical black hole undergoing accretion has already accumulated a large mass, while at low redshift, small black holes undergo accretion. Recent simulations have shown that such an anti-hierarchical growth is not contradictory to hierarchical structure formation (e.g. Fanidakis et al. 2012; Hirschmann et al. 2012; Enoki et al. 2014), and can explain the decrease of the peak of the luminosity function. Under the EB scenario, the peak of the luminosity function is linked to the Eddington luminosity, as brighter AGN remove their obscuration. If the black hole mass increases with redshift, so does the Eddington luminosity, causing the on-set of the high-luminosity decrease of the obscured fraction to occur at higher luminosities. The EB scenario is thus compatible with the observations, if the average black hole mass increases with redshift.

This model can affect both the small-scale, nuclear obscurer as well as the gas in the galactic vicinity of the black hole, and observations may be incapable of distinguishing the effects. For this reason, the illustration of this effect (EB) in Figure 6.3.2 might be placed in both categories.

Lets now consider the low-luminosity effect, which is a decrease of the obscured fraction towards low luminosities. The evolutionary scenario might also be capable of explaining the low luminosity increase of the obscuration (Hopkins et al. 2006a). If the observed column density is due to galactic streams, feeding the black hole simultaneously obscures it. One would then expect a simple correlation of the accretion luminosity and the obscuration, at least at luminosities below the Eddington limit. This effect (Gas Available, GA) is illustrated in the bottom left of Figure 6.3.2.

Under the GA scenario (bottom left of Figure 6.3.2), the accretion luminosity is simply correlated to the available gas. While observations have shown an overall reduction of the gas and dust content of galaxies over cosmic time (Tacconi et al. 2013), under the GA model this would increase the total number of AGN at high luminosities, but the luminosity-dependence of the obscured fraction would remain unaffected. In particular, the

GA model does not predict the observed increase of the turn-over luminosity as suggested by Figure 6.2.3.

The low-luminosity increase of the obscured fraction has been observed before in local samples (Burlon et al. 2011; Brightman & Nandra 2011b; Elitzur & Ho 2009). For these sources, it is more probable that the obscurer is nuclear, as many of these sources provide the opportunity to study both the galaxy and the infrared emission of the torus. One model is that the obscurer is caused by clumps in the disk wind (Elitzur & Shlosman 2006). This scenario was originally created to explain the vertical support for a cold dusty torus that would otherwise collapse (Krolik & Begelman 1988). In the disk wind model, the accretion disk is only capable of projecting clouds beyond a certain luminosity, predicting the absence of a clumpy torus in low-luminosity AGN. This scenario of cloud production (CP) is illustrated in the top left of Figure 6.3.2.

While the cloud production model reproduces the necessary increase of the obscured fraction, it does not predict any evolution over cosmic time when taken at face value. However, under this model, the critical luminosity below which the obscurer can not be sustained is strictly determined by the black hole mass (Elitzur & Shlosman 2006; Elitzur 2008). This model has received observational support in Elitzur & Ho (2009) who observe a black hole mass dependence of the luminosity dependence from a local survey. This model, taken together with the black hole mass evolution explained above, can thus explain the observed evolution.

To conclude, I have considered four effects (simple models) which are summarized in Table 6.2. Individually, they only partially explain the observations. However, the CP model, taken together with the EB scenario can explain the observations if the black hole mass function evolves with redshift. Under this view, the obscurer is vertically extended above a certain luminosity, and the torus disappears when the Eddington luminosity drives away the obscurer, unveiling the bright X-ray source for a (cosmically) brief time. Here it remains uncertain whether the obscuration removed is galactic, nuclear, or both. Both effects engage at a luminosity that is dependent on the black hole mass, which provides the cosmic evolution effect.

It is worth pointing out a consequence of the luminosity-dependence of the obscured fraction: Any study separating AGN by obscuration in a sample will bias the obscured sub-sample to lower intrinsic luminosities. If the systems under study scale in mass with luminosity and redshift as argued above, this may provide an explanation for some findings that are apparently obscuration-dependent, such as that obscured AGN show lower clustering (e.g. Allevato et al. 2014, but see also Georgakakis et al. 2014).

Future observations may shed light on the dependence of the galactic gas with the AGN X-ray luminosity. For instance, ALMA (Atacama Large Millimeter/Submillimeter Array, Beasley et al. 2006) measurements of gas fractions in host galaxies of unobscured and obscured AGN can be compared. If they do not differ, the GA model is ruled out. A similar approach by comparing luminous and faint AGN can also probe the predictions of the EB model, and establish whether the obscuration is nuclear. Under the EB model, gas motions are expected in bright AGN, which may be detectable with future X-ray spectroscopes on the ATHENA+ mission (Nandra et al. 2013). Furthermore, to establish

Table 6.2: Predictions of the discussed models.

Model ^a	Obscuration ^b	Low- L effect ^c	High- L effect ^d	Evolution ^e
RT	nuclear	✓		BHM ^f
CP	nuclear		✓	no
EB	nuclear or galactic		✓	BHM ^f
GA	nuclear	✓		no
RT+CP	nuclear	✓	✓	not for high-L
GA+EB	galactic	✓	✓	not for low-L
CP+EB	nuclear or both	✓	✓	BHM ^f

^a Model name as presented in Figure 6.3.2.

^b Whether the luminosity-dependent layer of Obscuration is associated with galaxy-scale obscuration (“galactic”) or the “torus” (“nuclear”).

^c Whether the model predicts a decrease of the fraction of obscured AGN towards low luminosities.

^d Whether the model predicts a decrease of the fraction of obscured AGN towards low luminosities.

^e Whether the model predicts any evolution of the effect with cosmic time.

^f The model scales with black hole mass. Evolution is predicted if the average accreting black hole mass changes over cosmic time.

that the evolutionary trend is caused by black hole mass evolution, black hole mass-matched samples of luminous AGN may be considered. If such a sample shows no redshift evolution in the luminosity-dependence of the obscured fraction, then the trend is due to black hole mass differences. The CANDELS fields with their high-resolution near-infrared imaging is the best candidate for such a study.

The models as presented here are a rudimentary description of the involved physical processes and require further refinement in their predictions and self-consistency, e.g. via numerical simulations. Among the challenges in this regard is that the nuclear obscuration can not be resolved in evolutionary models (e.g. Hopkins et al. 2006a; Somerville et al. 2008) and thus the torus is not yet treated self-consistently. Recently, it has become possible to self-consistently model the radiative and hydrodynamic processes that maintain the torus (e.g. Wada 2012). Further research is needed in terms of whether the torus can, by itself, reproduce the luminosity dependence of the obscured fraction. Also, it has to be clarified which processes maintain a cold, Compton-thick torus, with the geometric extent implied by a 35% Compton-thick fraction.

6.4 Conclusions

This work has combined deep and shallow, wide-area X-ray surveys conducted by Chandra and XMM-Newton to constrain the space density of X-ray selected AGN as a function of accretion luminosity, obscuring column density and redshift. An important feature of the analysis presented is the non-parametric method developed, which does not require any assumptions on the shape of the luminosity function or its evolution, allowing the data to drive the results. Furthermore, all sources of uncertainties are taken into account, allowing robust constraints on the evolution of unobscured and obscured AGN, including the most heavily obscured Compton-thick.

Obscured AGN, with $N_H > 10^{22} \text{ cm}^{-2}$, account for $77_{-5}^{+4\%}$ of the number density and $74_{-5}^{+4\%}$ of the luminosity density of the accretion SMBH population averaged over cosmic time. Compton-thick objects, with $N_H > 10^{24} \text{ cm}^{-2}$ account for approximately half the number and luminosity density of the obscured population, and $38_{-7}^{+8\%}$ of the total.

There is evidence that the space density of obscured, Compton-thin AGN evolves stronger than the unobscured or Compton-thick AGN. This is connected to the luminosity-dependent fraction of obscured AGN. At higher luminosities, fewer AGN are obscured. However, at higher redshift, this effect sets on at significantly higher luminosities. In the luminosity range used in this study, the fraction of obscured AGN increases from $75_{-4}^{+4\%}$ to $83_{-3}^{+3\%}$ at $z > 2.25$. This is due to only a small luminosity range around $L_* \approx 10^{44} \text{ erg/s}$ which changes its obscured fraction. In particular, the space density evolves fastest around $N_H \approx 10^{23} \text{ cm}^{-2}$. In contrast the fraction of Compton-thick AGN relative to the total population is consistent with being constant at $\approx 35\%$ independent of redshift and accretion luminosity. The contribution to the luminosity density by Compton-thick AGN is $40_{-6}^{+6\%}$. The robust determination of a large fraction of Compton-thick AGN consolidates AGN X-ray surveys and studies of the cosmic X-ray background. It also implies a physically

extended nuclear obscurer.

Also, the fraction of moderately obscured AGN decreases not only to high but also to low luminosities, in qualitative agreement with findings of local surveys. I discussed the observed trends of the obscured fraction with L_X and z in the context of models that either assign obscuration to the torus or galaxy-AGN co-evolution effects, with varying luminosity-dependent effects. Both classes of models can qualitatively explain the results but require that SMBH evolve in a downsizing manner, i.e. larger black holes form earlier in the Universe.

7 Accretion history of the Universe

Supermassive Black Holes (SMBH) are abundant in the local universe – every nearby massive galaxy harbours one (Richstone et al. 1998; Kormendy & Ho 2013). The majority of the growth of these SMBH must have been through accretion processes (Soltan 1982; Merloni & Heinz 2008). The accretion disk around black holes release radiation power (luminosity) proportional to the accreted mass, to release the angular momentum and potential energy of the incoming matter (Shakura & Sunyaev 1973). Studying the luminosity output of the AGN population thus gives access to the mass accreted into black holes over cosmic time. Due to the large fraction of obscured AGN (e.g. Maiolino & Rieke 1995; Risaliti et al. 1999, and references in the previous chapter), the total luminosity output of unobscured AGN (as selected by optical/UV-selected samples) may not be sufficient to explain the large masses known to reside in black holes of nearby galaxies. Indeed, a quantitative comparison by Marconi et al. (2004) consolidates the two observables by requiring that the majority of the growth must occur in a hidden phase (obscured AGN). The contribution of Compton-thick AGN to the luminosity function in particular, which was assumed in their calculation, has remained a subject of debate. In this work, a hard (2 – 10 keV) X-ray selection is used, which is less sensitive to obscuration, and makes it possible to quantify the selection bias due to line-of-sight obscuration. This Chapter, based on the non-parametric luminosity framework developed in Chapter 6, investigates the total accretion luminosity output of the AGN population, its evolution, and the importance of obscured AGN in this context. Understanding the total luminosity output of the AGN population requires understanding the total number of AGN and the shape of their luminosity function, as well as any changes over cosmic time of either (evolution with redshift), which is discussed in Section 7.2. In Section 7.3, the X-ray luminosity is converted into mass accretion rate to compare the luminosity output to the local black hole mass density.

7.1 Shape of the X-ray Luminosity function

7.1.1 Introduction

The intrinsic luminosity distribution of AGN can be modeled using a double powerlaw distribution (see Figure 7.1.1 for an illustration). In the most commonly adopted rendition, named “bending powerlaw” here, this luminosity function can be formulated as

$$\frac{d\Phi(L)}{d\log L} = \frac{\phi_0}{\left(\frac{L}{L_*}\right)^\alpha + \left(\frac{L}{L_*}\right)^\beta}, \quad (7.1.1)$$

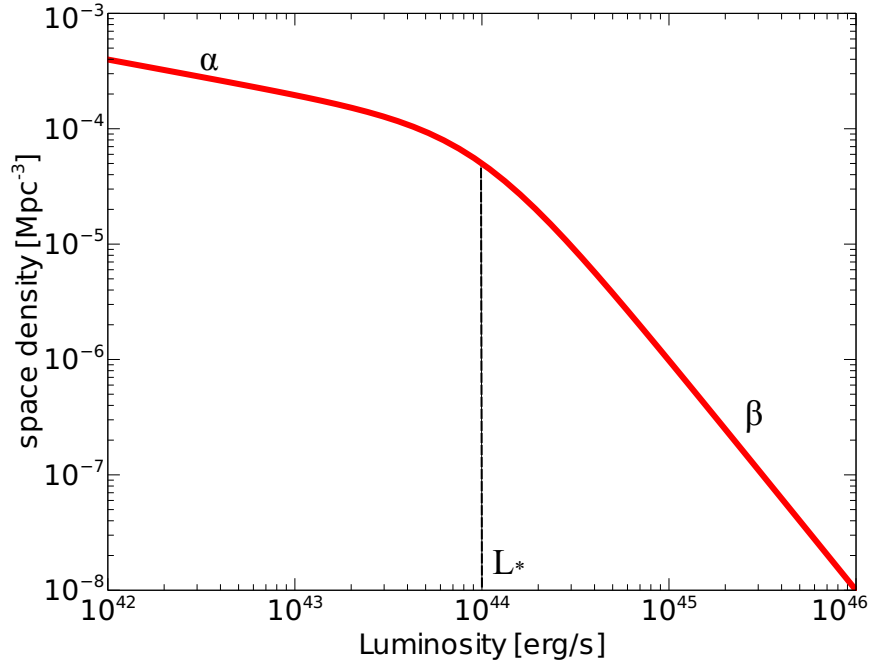


Figure 7.1.1: Illustration of a bending powerlaw distribution. The number density of AGN at faint luminosities is substantially higher than at bright luminosities. Additionally, at a certain luminosity (L_*), the powerlaw distribution changes its slope.

where the distribution, illustrated in Figure 7.1.1, changes behaviour at the break luminosity L_* , which occurs in hard X-rays at $L_* \approx 10^{44}$ erg/s (e.g. Marshall et al. 1983; Ueda et al. 2003; La Franca et al. 2005; Aird et al. 2008, 2010; Ueda et al. 2014). This function is relatively flat ($\alpha \approx 0 - 0.5$) at low luminosities below L_* (the faint end). However, at high luminosities (the bright end) a steep decrease ($\beta \approx 2 - 3$) in the number density is observed. The normalisation at L_* is denoted by ϕ_0 .

The evolutionary behaviour of the luminosity function has a long history of debate. Initially, two models of evolution have been proposed: (a) Pure Density Evolution (PDE, Schmidt 1968), where the shape remains constant, but the normalisation declines from towards $z = 0$ and (b) Pure Luminosity Evolution (PLE, Mathez 1976), where the luminosity of all sources declines which also produces a decline in the overall number of AGN at each luminosity. Schmidt & Green (1983) concluded, based on an optical (B-band) selected AGN sample that neither models are appropriate and introduced the Luminosity Dependent Density Evolution model (LDDE). This model has since received support from a number of studies analysing the soft (Miyaji et al. 2000; Hasinger et al. 2005) and hard band (Silverman et al. 2008; Yencho et al. 2009; Aird et al. 2010), as well as intrinsic (absorption-corrected) hard band (Ueda et al. 2003; La Franca et al. 2005; Ebrero et al. 2009; Ueda et al. 2014) luminosity function, which established that LDDE describes the data substantially better than PDE or PLE. The evolutionary behaviour of the LDDE model modifies the normalisation, break luminosity and the faint end slope. Whether a

modification of the faint end slope is required has been called into question by Aird et al. (2010), who contribute significant methodological improvements in terms of accounting for X-ray sources without optical counterparts, photometric redshift uncertainties and employing Bayesian model comparison. They present an alternative model, Luminosity and Density Evolution (LADE) wherein the shape of the luminosity function does not evolve. In their study, they find no evidence for LDDE over LADE if uncertain counterparts and problematic redshifts are excluded. This may indicate that systematic uncertainties drive the preference towards LDDE in studies not accounting for multiple counterparts or the systematic uncertainties of photometric redshifts, even when a large data sample is used (e.g. the most recent work by Ueda et al. 2014).

The shape of the luminosity function and its evolution has to be interpreted in context of large-scale structure assembly and galactic co-evolution processes (“feedback”). However, if only the normalisation of the luminosity function increases (PDE), then it may be that only the number of host galaxies increases. Also, if only the break luminosity evolves (PLE), it may be that this behaviour follows the break of the galactic stellar mass function (M_*) through scaling laws between host galaxy properties and SMBH. Such scaling laws are known in the local Universe (Magorrian et al. 1998; Ferrarese & Merritt 2000; Gebhardt et al. 2000; Tremaine et al. 2002; Häring & Rix 2004; Ferrarese & Ford 2005; Kormendy & Ho 2013), but whether they hold to high redshift is unclear. A change in the shape of the luminosity function would require more severe alterations of the co-evolution picture (feedback processes, triggering mechanisms).

7.2 Observed luminosity function

In the previous chapter I have reconstructed the distribution of AGN in luminosity, redshift and obscuration. There I have investigated the relative importance of e.g. unobscured and obscured AGN. In this chapter, I focus the luminosity distribution of AGN of all obscurations. To this end, Figure 7.2.1 presents the total luminosity function from a number of redshift intervals. Here, the non-parametric total X-ray luminosity function has been integrated along the N_H axis for this plot. The results from both smoothness priors (see Chapter 6) have been combined. Typically, the constant-value prior provides the lower limits at the faint end and upper limits at the bright end by extrapolating from the densities at intermediate luminosities. In contrast, the upper limits at the faint end and the lower limits at the bright end are produced by the constant-slope prior, which tends to invoke a steep slope. Large deviations (uncertainties) are visible at the faint end, in particular at high redshifts $z > 2$, as the corresponding low X-ray fluxes begin to fall below the sensitivity limit of the survey, which limits the number of detected sources. At all redshift intervals, the luminosity function appears to be consistent with a bending powerlaw shape (as shown in the independently fitted dotted line), although with varying normalisation and break luminosity. Because the data constraints are poor at the bright end, the bright-end slope was fixed in the fitting to $\beta = 2.3$ (Aird et al. 2010), which is very close to the results based on the constant-slope prior (lower limit of the error

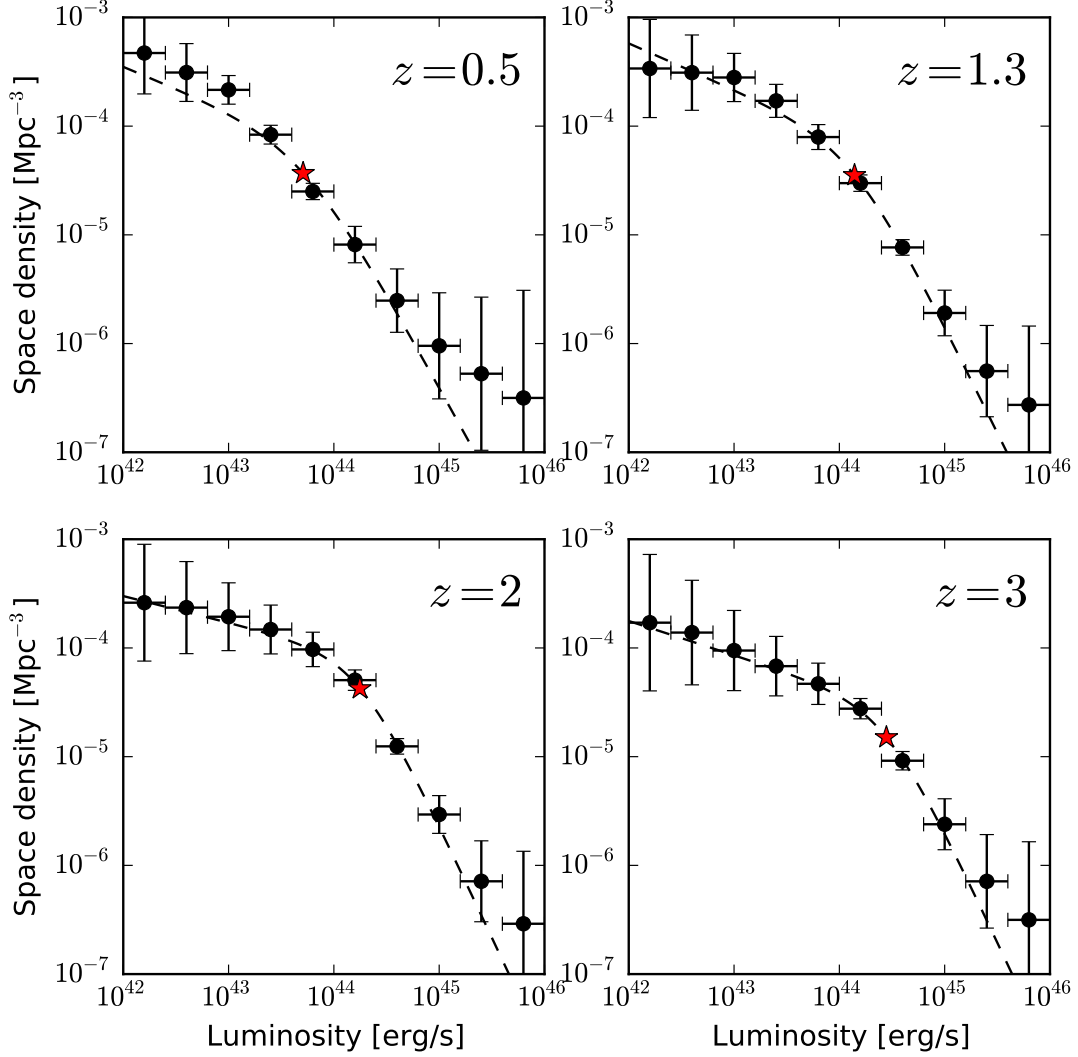


Figure 7.2.1: The 2-10 keV luminosity function at various redshift bins. Black data points show the non-parametric reconstruction as a function of luminosity, using the combined posterior distributions from both smoothness priors which provide upper and lower limits (see Chapter 6). The dashed lines indicate a bending powerlaw fit to each redshift shell independently, with the bright-end slope fixed to $\beta = 2.3$, to follow the results based on the constant-slope prior at the bright end where our constraints are poor. The red star indicates the break luminosity L_* and its normalisation. The bending powerlaw provides a good fit to the observed luminosity function.

bars). It should be kept in mind that the error bars in Figure 7.2.1 are not independent measurements but are connected by a smoothness assumption. As such, they reflect the uncertainty in the true space density rather than the scatter of an estimator.

7.2.1 Evolution of the luminosity function

The results of the fitting of the bending powerlaw in each redshift shell is shown in Figure 7.2.2. Each fit was run 400 times from random start values. The distribution of final fit values is recorded as 1-sigma error bars in Figure 7.2.2 for the normalisation, break luminosity L_* and faint-end slope α . The normalisation shows a broad plateau around $z = 2$. This is related to the evolution of the total number of accreting AGN and will be discussed below in the context of the luminosity density in Section 7.3.2. The most notable evolutionary behaviour is the increase of the break luminosity, L_* , with increasing redshift. At low redshifts ($z = 0.5$), the break luminosity is $L_* \approx 10^{43.3}\text{erg/s}$ and rises to $L_* \approx 10^{44.2}\text{erg/s}$ by $z = 2$, where it appears to remain constant. No evolution of the faint-end slope can be detected. In Figure 7.2.2, the slope at low luminosities is consistent with zero or small positive values $\alpha = 0 - 0.5$, with no evidence for evolution.

It is now also possible to estimate the total number of AGN by integrating the luminosity function over a range of luminosity and redshift (weighted by the comoving volume). The number of AGN at $z < 4$ is $2.2^{+3.3}_{-0.9} \times 10^8$ (with 90% probability, for $L > 10^{42}\text{erg/s}$). The bottom panel of Figure 7.2.2 depicts the evolution of the comoving number density over redshift. Overall, the number density is approximately constant. However, at redshifts above $z = 2$ there is some evidence for a decline in the total number of AGN within the integration window of $L = 10^{42-46}\text{erg/s}$.

7.2.2 Comparison with previous works

Figure 7.2.3 shows the non-parametric X-ray luminosity function, integrated along the N_H axis in small redshift intervals. Black error bars show the estimation of densities from the non-parametric reconstruction. Again, these error bars are derived by combining the posterior distributions based on the constant-value prior and the constant-slope prior. These deviate in behaviour outside the strongest data constraints, both at the highest and lowest luminosities. As before, the constant-value prior tends to keep the density constant and thus defines the lower limits at the faint end and the upper limits at the bright end. In contrast, the constant-slope prior tends to keep the slope of the powerlaw constant and declines strongly with luminosity. This defines the upper limits for the faint end and the lower limits at the bright end. Some of the strongest deviations (and thus uncertainties) are visible at the bright end at low redshift ($z < 0.3$), where the sample is small due to the limited cosmological volume. At high redshift ($z > 2$), the faint end becomes uncertain as the X-ray sensitivity limits the number of detected sources.

The works of Ueda et al. (2003) (blue line), Aird et al. (2010) (orange line) and Ueda et al. (2014) (red line) are overplotted in Figure 7.2.3. For comparison between the panels, the $z = 0.1$ luminosity function of Ueda et al. (2014) is repeated in every panel (dotted

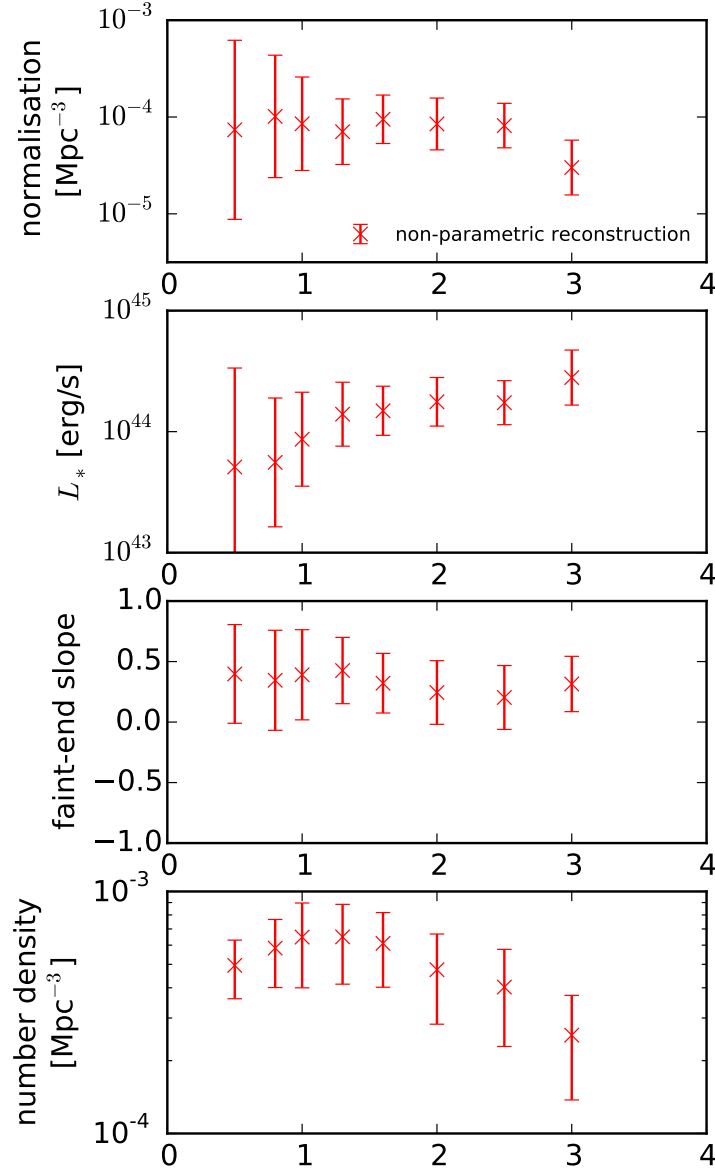


Figure 7.2.2: The evolution of luminosity function parameters. Data points represent bending powerlaw fits to the non-parametric reconstruction, as shown in Figure 7.2.1. The top three panel show the normalisation at L_* (*top*), the break luminosity L_* (*centre*) and the faint-end slope α (*bottom*). The bright-end slope was fixed to $\beta = 2.3$. Finally, the *bottom panel* shows the total number density of AGN. Here, the luminosity function was integrated within $L = 10^{42-46}$ erg/s.

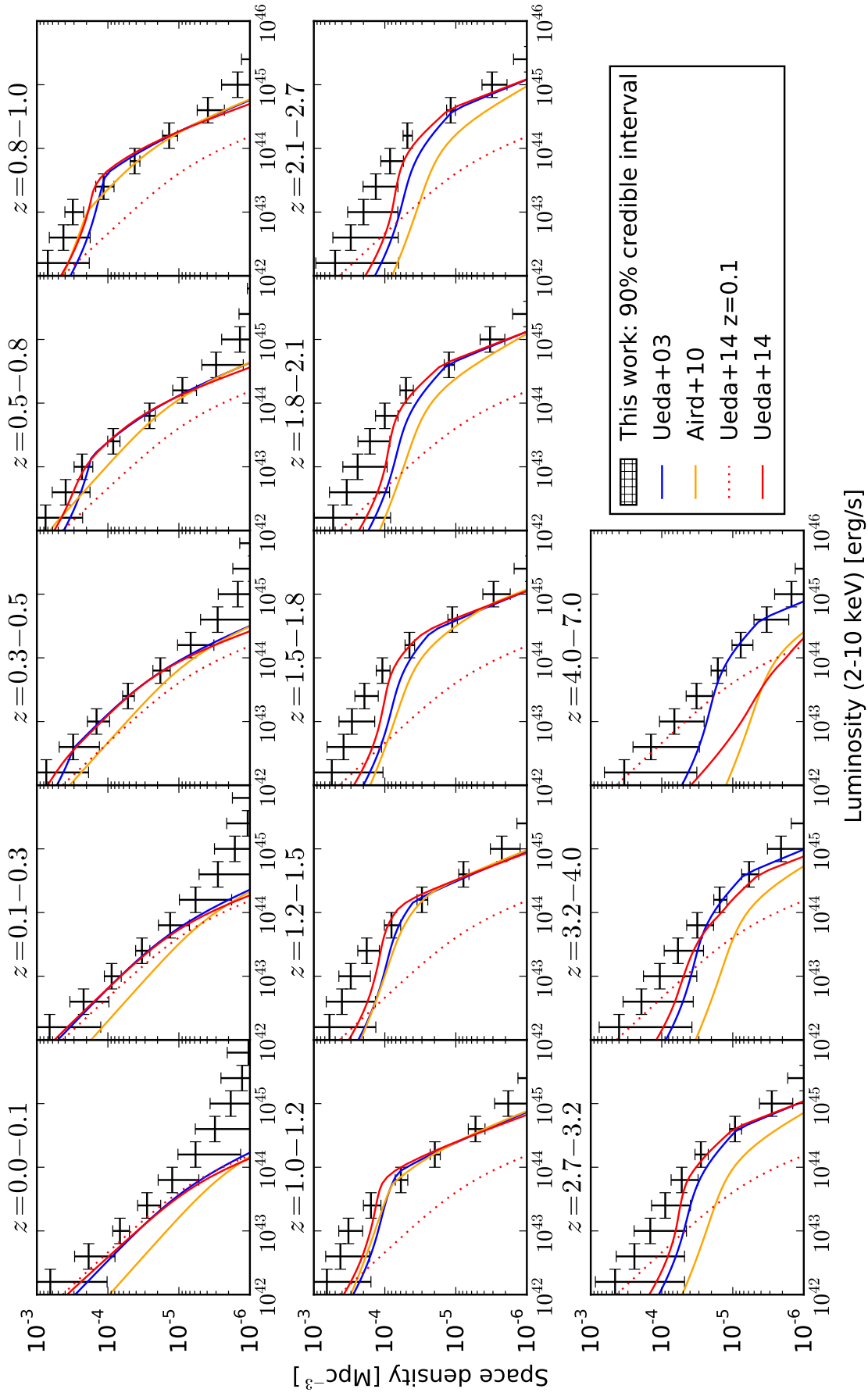


Figure 7.2.3: Total luminosity function. Each panel depicts a redshift shell. The data points plot the non-parametric luminosity function reconstruction, by combining the results from both smoothness priors. These appear consistent with a broken powerlaw at all redshifts. For comparison, the parametric fits of Ueda et al. (2003), Aird et al. (2010) and Ueda et al. (2014) are shown in blue, orange and red respectively. For comparison across panels, the Ueda et al. (2014) fit from $z = 0.1$ is kept across all panels as a dotted red line.

red). The non-parametric reconstruction shows close similarities to previous parametric determinations in general shape, normalisation and evolution with redshift of the X-ray luminosity function (XLF). The overall shape of the luminosity function is a double power-law with a break or bend at a characteristic luminosity (L_*), the value of which increases with redshift. As found in previous studies, the space density shows a rapid evolution up to around $z \sim 1$ at all luminosities, being most prominent at high luminosities due to the positive evolution of L_* . The positive evolution appears to continue up to $z \sim 3$ (as in Aird et al. 2010), above which the population starts to decline. In the comparison of Figure 7.2.3 it is important to keep in mind that the error bars are not independent. Figure 7.2.3 also compares to a recent comprehensive study of the XLF by Ueda et al. (2014). Their is in overall agreement with the results derived by the constant-value prior at low luminosities and those obtained by the constant-slope prior at high luminosities. It is worth emphasising that in the non-parametric approach these features are imposed by the data and not by any assumptions about the functional form of the X-ray luminosity function.

While the general behaviour of the population is similar, Figure 7.2.3 also shows important differences compared to some previous parametric studies, specifically that of Ueda et al. (2014). At the highest redshift bin ($z = 4 - 7$), the space density drops sharply towards high redshifts in their XLF reconstruction (see also Civano et al. 2011; Kalfountzou et al. 2014; Vito et al. 2014). A minority of the sources used in this study is at $z > 4$, and only a small fraction of these sources has spectroscopically confirmed redshifts (see Figure 4.8.1 on page 65). However, the non-parametric reconstruction remains, as suggested by the priors, at space densities comparable to the previous redshift bin $z = 3 - 4$. In the data used in this work, there seems to be no evidence of a steep decline with redshift. The difference may be due to the large uncertainties in the redshift estimates used in this work, and the coarse binning in redshift.

At intermediate redshifts, the XLF of Ueda et al. (2014) shows a sharp flattening below a luminosity of around 10^{44} erg/s, after which it steepens again at the lowest luminosities probed by the study. This behaviour is most apparent in Figure 7.2.3 at redshifts $z \sim 1 - 2$. It is only easy to parameterise this shape using the Luminosity Dependent Density Evolution model (e.g. Ueda et al. 2003; Hasinger et al. 2005; Silverman et al. 2008). The analysis performed here does not support such a behaviour. This is most apparent in the redshift range $z = 0.8 - 2.1$, where the non-parametric field analysis requires a significantly larger space density of AGN in the critical luminosity range of 10^{43-44} erg/s. This luminosity range is close to the survey flux limit of the surveys used in this work, but also others, and thus missing counterparts and mis-identification can play an important role. At these redshifts it is also difficult to estimate correct redshifts, due to the absence of emission line features. Previous works have typically neglected detected X-ray sources with missing optical/infrared counterparts and the uncertainties in redshift estimates. In this work, these uncertainties are taken into account in the form of redshift probability distributions. The difference in methodology may explain the “missing” density in other works as compared to this work.

In Section 7.4 I continue by interpreting the shape of the luminosity function and dis-

cussing it in context of the literature. In the following section I link the luminosity output of the AGN population to the accreted mass and discuss the growth of black holes over cosmic time.

7.3 The accretion history of the Universe

7.3.1 Inferring infalling matter from X-ray radiation

AGNs are powered by accretion onto black holes, with their emission originating in the partial release of potential energy in accretion disks (see Section 3.2.2). Through the relation of radiation and accreted matter it is possible to study the accretion history of SMBHs (their build-up over cosmic time), which is investigated in this section. A feature of this work is that robust constraints on Compton-thick AGN are included in the calculation of the accretion density.

The problem of converting between X-ray luminosity and mass accretion rate is separated here into two steps, which are illustrated in Figure 7.3.1. First, the luminosity released as a function of mass accretion rate has to be defined (see Section 3.2.1). Here, a fixed radiative efficiency of $\mu = 10\%$ is adopted (see Section 7.4.2 for a discussion).

Secondly, as this study measured X-ray luminosities, the bolometric luminosity has to be converted into the luminosity output in the $2 - 10\text{keV}$ band. Many works have investigated the bolometric luminosity distribution of AGN from observations of bright, nearby objects (e.g. Elvis et al. 1994; Richards et al. 2006a) and from a population perspective (Hopkins et al. 2007). Recent studies find that the bolometric corrections are not a constant, but vary with luminosity (e.g. Marconi et al. 2004; Hopkins et al. 2007; Lusso et al. 2012). However, as this work uses intrinsic (unabsorbed) X-ray luminosities, it is important in this step not to double-count the infrared dust emission of the torus, which is re-radiated UV and X-ray emission. Such bolometric corrections have been measured (e.g. Marconi et al. 2004; Lusso et al. 2012). Here we adopt the conversion formula from Marconi et al. (2004), who use the approximation

$$\log \left[\frac{L_{bol}}{L(2 - 10\text{keV})} \right] = 1.54 + 0.24\mathcal{L} + 0.012\mathcal{L}^2 - 0.0015\mathcal{L}^3, \quad (7.3.1)$$

where $\mathcal{L} := (\log \frac{L_{bol}}{L_{\odot}} - 12)$ and L_{\odot} is the bolometric luminosity of the sun. In brief, the fraction of energy emitted in the $2 - 10\text{keV}$ band is of the order of 10% of the bolometric luminosity. However, at the brightest bolometric luminosities ($L \approx 10^{13}L_{\odot} \approx 3.83 \times 10^{46}\text{erg/s}$), this fraction decreases to $\approx 2\%$. This conversion is shown in the top panel of Figure 7.3.1.

The relationship of Equations 3.2.1 and 7.3.1 allow, for a given black hole mass, to infer for a given X-ray luminosity the mass accretion rate onto the black hole and thus its instantaneous growth rate. Using the luminosity function ϕ , the integration over luminosities, $\dot{M}_{total} = \int \phi(L_X) \cdot \dot{M}(L_{bol}(L_X)) d\log L_X$, describes the growth rate of all black

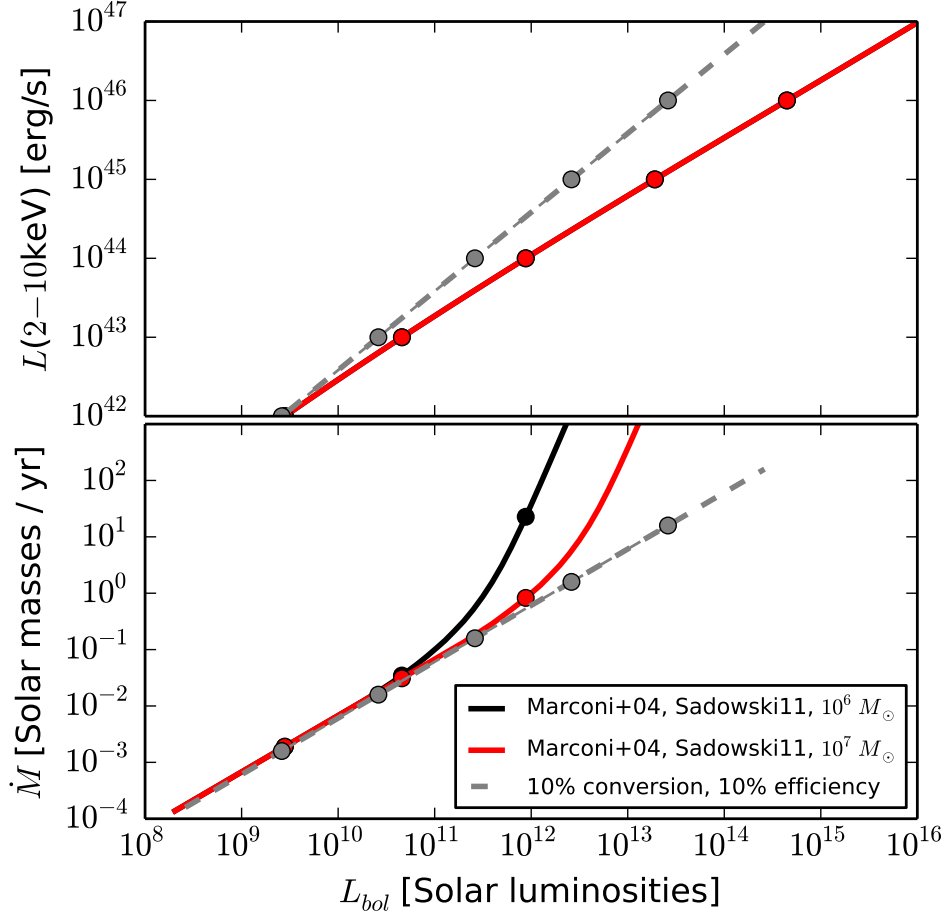


Figure 7.3.1: Conversion between X-ray luminosity in the 2 – 10keV band (erg/s), bolometric luminosity (in L_\odot) and accretion rate (in M_\odot per year). *Top panel:* The fraction emitted in the X-ray band is 10% at low luminosities (gray dotted line), but is observed to decrease towards higher luminosities (red line, from Marconi et al. 2004). *Bottom panel:* The accretion disk converts, to first order, approximately 10% of the rest mass of the accreted mass into radiation (gray dotted line). Simulations show that this fraction depends on spin and black hole mass. The luminosity conversions for $10^6 M_\odot$ (red line) and $10^7 M_\odot$ (black line) non-rotating black holes are shown (from Sadowski 2011).

holes at a given time. Integrating \dot{M}_{total} over the history of the Universe yields the amount of matter trapped in black holes (black hole mass density, BHMD) at $z = 0$.

7.3.2 The importance of accretion in obscured black holes

Before applying any conversions, I compute the total X-ray luminosity density in the 2–10keV X-ray band. The luminosity density is the integration of the luminosity function as $\int L_X \cdot \phi d\log L_X$ at a certain redshift. This is shown in the top of Figure 7.3.2 (grey). Here, I again use $L > 10^{43.2}\text{erg/s}$, which captures essentially all of the luminosity density and I focus on the redshift range $z = 0.5 - 4$ where the data coverage of this sample is best. The luminosity density shows a broad peak in the $z = 1 - 3$ range, with a decline to both high and low redshifts.

Many studies determine accretion luminosity evolution, however residual uncertainties remain in these determinations, because of the potentially large population of Compton-thick AGN in the Universe. The work of Chapter 6 has placed constraints on the Compton-thick AGN population. To highlight, Figure 7.3.2 also probes the contribution of AGN split by their obscuration (Unobscured, Compton-thin obscured, Compton-thick). One of the most striking results is that Compton-thick AGN and obscured Compton-thin AGN contribute in equal parts to the luminosity density in the $z = 1 - 3$ range. These obscured AGN contribute the majority ($74^{+4}_{-5}\%$), while unobscured AGN only play a minor role. The shapes in evolution are fairly similar. However, the evolution of unobscured AGN is weaker compared to the Compton-thin obscured AGN when considering the $z \sim 1 - 1.5$ range. This effect is related to the increase in the obscured fraction with redshift (see Section 6.2.3).

Figure 7.3.3 shows the mass accretion rate density of AGN over cosmic time, i.e. the mass trapped into black holes per gigayear (Gyr) per cubic megaparsec. Overall, the black hole mass density grows by a rate of $\approx 10^5 M_\odot/\text{Mpc}/\text{Gyr}$. Tentatively, the reconstruction has a similar shape as the luminosity density in Figure 7.3.2. However, the bolometric conversion of Equation 7.3.1 was used here, which gives additional weight to the poorly constrained bright end, increasing the uncertainties.

The mass accretion rate density can now be integrated over cosmic time. Figure 7.3.4 shows the mass per comoving volume locked into black hole masses at various stages in cosmic time. For comparison, local estimates of the black-hole mass density based on the black hole masses of nearby galaxies are shown as error bars (Shankar et al. 2004; Marconi et al. 2004). The luminosity output of unobscured AGN (red shades) can not reproduce the black hole mass density found in the local universe ($z = 0$). Only when obscured AGN are also considered, the luminosity output over the cosmic history matches the black hole mass density. In this plot, the upper end of the uncertainties stem from the uncertain bright end. Such bright AGN are rare, but contribute substantially in terms of total mass accreted. If the luminosity function is a powerlaw shape at the brightest luminosities, as assumed in parametric models, the lower end of the uncertainties is more relevant, and shows excellent agreement with local estimates. The upper end of the uncertainties is driven by the constant-value prior. For the bright end to be better constrained using the

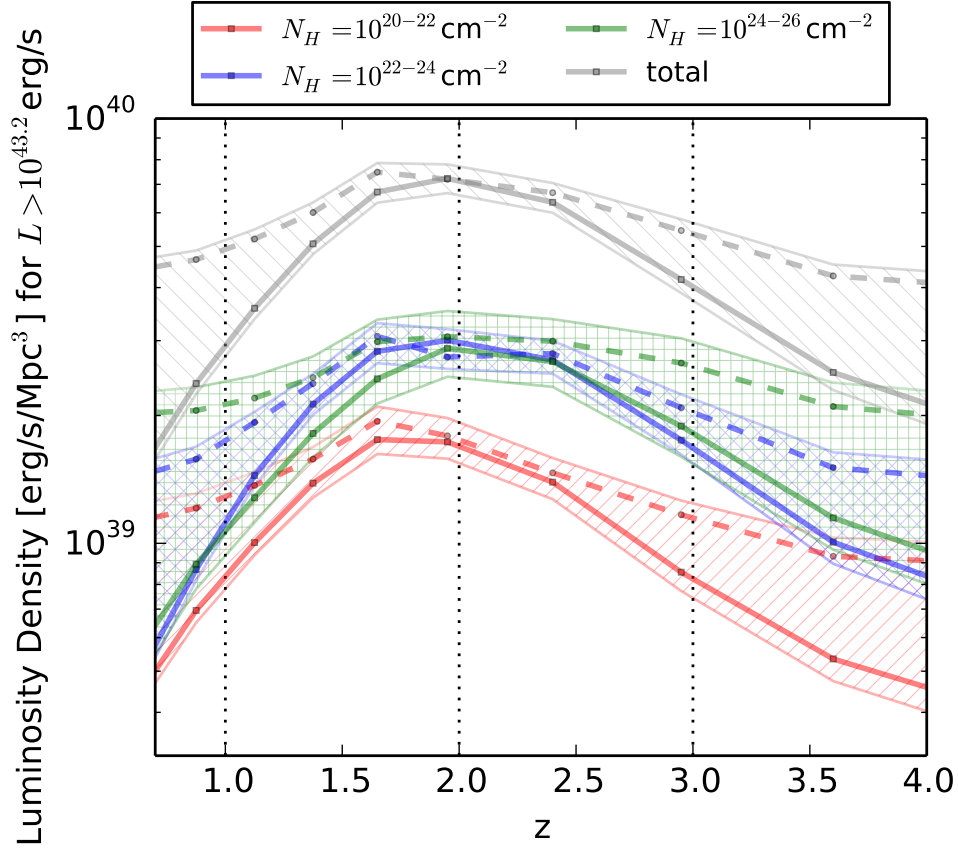


Figure 7.3.2: Evolution of the 2 – 10keV X-ray luminosity density of AGN with $L_X > 10^{43.2}$ erg/s, for various column densities. The luminosity output of AGN experiences a rise and fall in density in the $z = 1-3.5$ range (total as top gray shaded region). The strongest contribution to the luminosity density is due to obscured, Compton-thin (blue shaded region) and Compton-thick AGN (green shaded region), which contribute in equal parts to the luminosity. In contrast, the emission from unobscured AGN (red shaded region, bottom) is smaller.

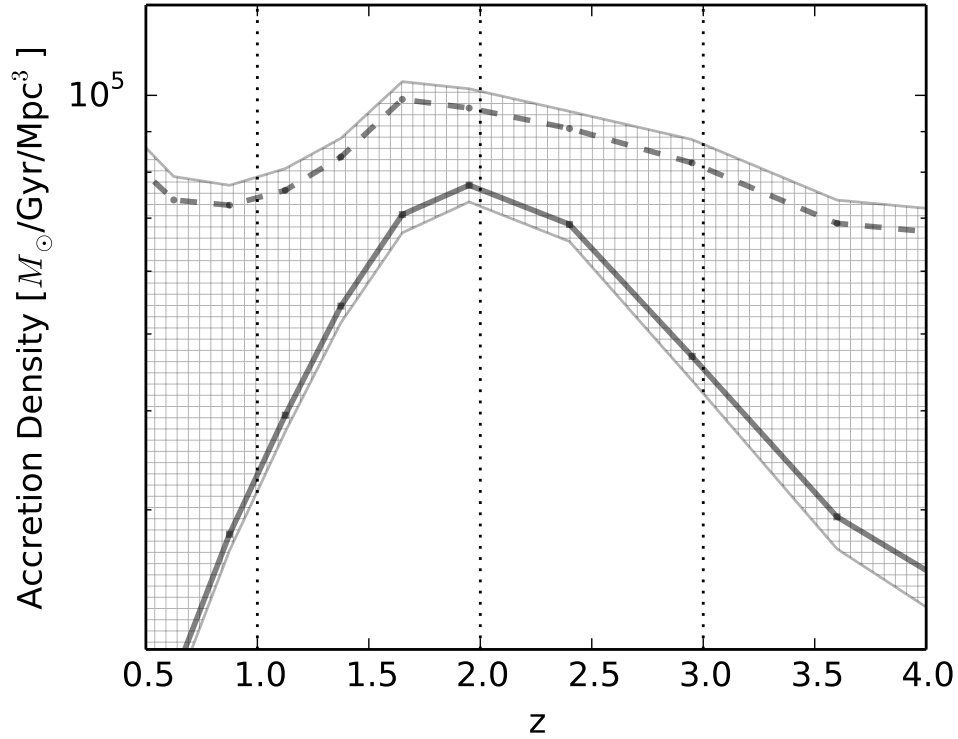


Figure 7.3.3: Evolution of the accretion density of AGN. The total bolometric luminosity output of AGN is converted to accretion rates by assuming 10% of the rest mass energy of accreted matter is radiated away. The dashed line corresponds to the result of the constant-value prior, while the solid line corresponds to the result of the constant-slope prior. The constant-value prior drives the result to systematically higher values because of its high density estimate at the bright end function (see Figure 7.2.3). A bending powerlaw model, as often assumed, would yield a result comparable to the constant-slope prior.

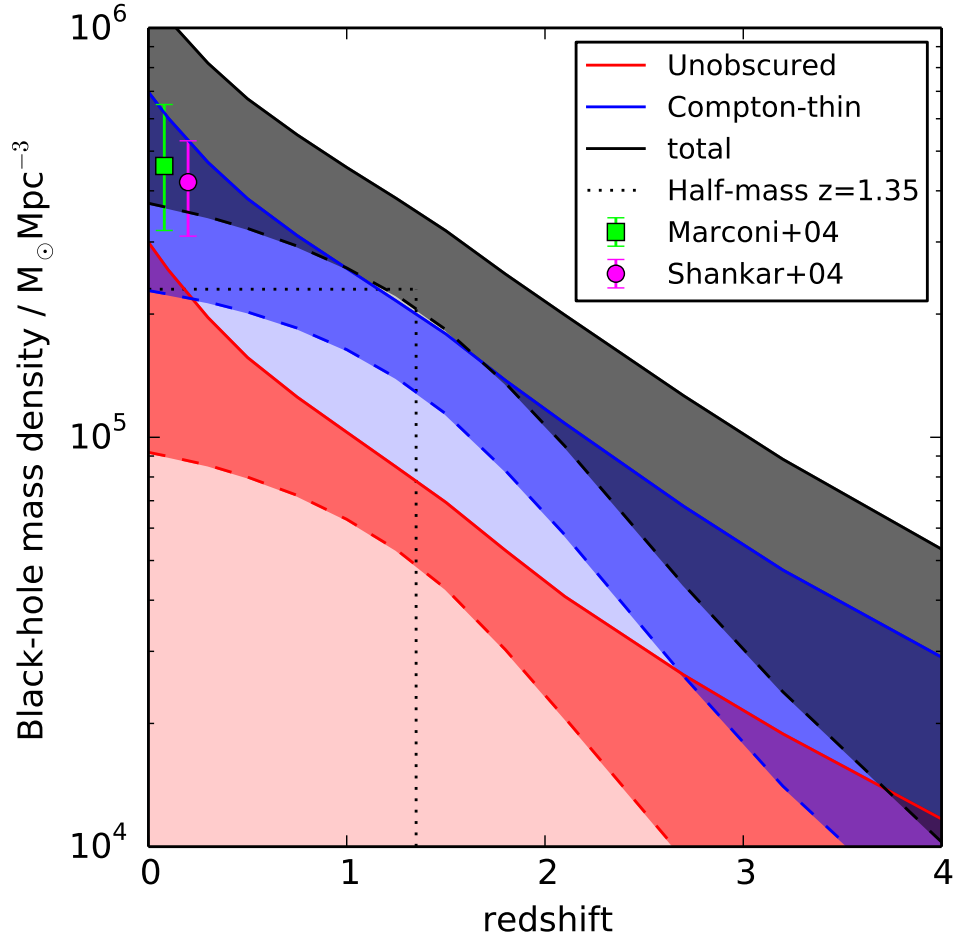


Figure 7.3.4: Black hole mass density. The mass accretion rate density (Figure 7.3.3) was integrated over cosmic time to estimate the mass trapped in black holes per cubic megaparsec. The red shades show the integration using only unobscured AGN ($N_H < 10^{22} \text{cm}^{-2}$), with the dark region representing the uncertainty based on 90% probability regions. The blue shades show the integration of all Compton-thin AGN ($N_H < 10^{24} \text{cm}^{-2}$), while black shows the total AGN population. For comparison, local estimates of the black-hole mass density based on the black hole masses of nearby galaxies are shown as error bars (Shankar et al. 2004; Marconi et al. 2004). While uncertainties are large, the luminosity output of the total AGN population can reproduce the local black hole mass density, but only if obscured AGN are included.

non-parametric method, more data from large area surveys is needed.

7.4 Discussion

7.4.1 The shape and evolution of the luminosity function

The non-parametric method unveils the X-ray luminosity function without assuming its shape and evolution beforehand. Overall, the shape of the XLF can be described as a bending powerlaw at all redshifts (see Figure 7.2.1). Figure 7.2.3 compares the results of this work to those from previous studies, which have adopted parametric models. While the overall shape is similar, there are some differences. The Luminosity Dependent Density Evolution (LDDE) model of Ueda et al. (2014) produces a more complicated behaviour with an inflection point at intermediate luminosities ($L \approx 10^{43-44}$ erg/s). It has been suggested that the preference for the LDDE model, which is the only parameterisation producing such shapes, is driven by incompleteness in luminosities near the detection sensitivity limit (Aird et al. 2010). In this work, no evidence for such shapes is found. In all redshift shells, the non-parametric method appears consistent with a broken or bending powerlaw model (see also Figure 7.2.1).

Both the break luminosity and the normalisation evolve over cosmic time (as opposed to pure luminosity evolution or pure density evolution models). The centre panel in Figure 7.2.2 shows that the break of the luminosity function, L_* , decreases over cosmic time, in particular at $z < 2$. If the break is associated with the scale of the system – e.g. related via scaling relationships to the break of the galaxy stellar mass function (M_*) or through feedback limitations such as the Eddington luminosity – then the typical accreting black hole mass decreases with redshift: Massive black holes tend to grow early, while small black holes tend to dominate accretion at low redshifts. This anti-hierarchical growth is often termed “down-sizing” (e.g. Merloni 2004; Hasinger et al. 2005). Although counter-intuitive, it can be reproduced in cosmological simulations (e.g. Fanidakis et al. 2012; Enoki et al. 2014). In the simulation of Fanidakis et al. (2012) for instance, the co-evolution of AGN and host galaxies creates down-sizing via the relative importance of two different accretion channels. One channel is the accretion of hot gas from the halo, which is important throughout cosmic time to produce low-luminosity AGN (“radio mode”, Croton et al. 2006). Disk instabilities and minor mergers cause star bursts and accretion of cold gas which provides another important channel, in particular at $z = 1 - 3$ (peak of the cosmic star formation) for feeding massive black holes and causing brightest luminosities (“quasar mode”). When this channel loses importance towards low redshift, bright AGN disappear, and the peak of the luminosity distribution shifts to lower values.

The total luminosity output of the AGN population shows a broad plateau at $z = 1 - 3$ in Figure 7.3.2. A decline is apparent towards lower redshifts. This can be attributed to the evolution of the break luminosity (luminosity evolution) while the number density remains approximately constant (see Figure 7.2.2). Additionally, above $z = 3$, a decline in the total number density of AGN (density evolution) can be observed in Figure 7.2.2. An increase

in the total number of AGN over time requires either the formation of new SMBHs or the triggering of previously “dormant” black holes. This effect co-incides with the evolution of the star formation rate and suggests a relation between activation of accretion onto super-massive black holes and star formation (Merloni & Heinz 2013). Perhaps the simplest mechanism is a shared gas reservoir that feeds both processes. Under this hypothesis, the incidence of accretion onto black holes should follow the star formation in lockstep. Simple recipes have been shown to connect the AGN luminosity function and star formation history (Mullaney et al. 2012; Hickox et al. 2014). While Rosario et al. (2011) finds no evidence for a correlation between specific star formation rate and AGN activity, Rosario et al. (2013) shows that AGN live preferentially in star forming galaxies. They conclude that AGN are present whenever galaxies have some minimal star formation, but beyond that, the amount of accretion luminosity and star formation are uncorrelated. The expected correlation may exist on long timescales (100 Myrs), but is washed out by the strong variability of AGN (Hickox et al. 2014). Further circumstantial evidence that increased availability of gas boosts AGN activity comes from the observed increase of the molecular gas fraction towards $z = 1 - 2$ (Saintonge et al. 2011a,b; Tacconi et al. 2013).

The data used here lack strong constraints at the bright end of the luminosity function, prohibiting us from studying the evolution of the bright end slope. However, the faint-end slope is fairly well constrained to $\alpha = 0 - 0.5$ (third panel in Figure 7.2.2). No evolution of the slope is detected. In the model of Hopkins et al. (2005b), the slopes of the luminosity function can be related to the time spent above and below the Eddington-luminosity. In Hopkins et al. (2006b) the authors argue that a flattening of the faint-slope can be interpreted directly as a change in the behaviour of AGN. In their model, after a luminous, post-merger phase of the host galaxy, the brightly ignited AGN declines exponentially in luminosity and migrates towards the faint end of the luminosity function. The speed of this transition is found, via simulations, to be dependent on the peak luminosity attained. The faint-end slope then is a proxy for measuring the typical maximum luminosities, as brighter peak luminosities cause a slower transition, and thus a flatter slope. A change in the faint-end slope would thus correspond to a variation in the peak luminosities typically attained. Some works in optical wavelengths have claimed an evolution in the slopes (e.g. Richards et al. 2006b; Croom et al. 2009). However, the separation of intrinsic evolution, contamination and selection biases is substantially more complicated there than in the hard X-ray window. All X-ray luminosity function works based on popular LDDE model also invoke an evolution of the faint-end slope due to the parameterisation of that model. Aird et al. 2010 has argued that this feature is unnecessary and proposed a new model in which the faint-end slope is constant. Also the non-parametric approach of this work does require a faint-end slope evolution.

7.4.2 Accretion history of the Universe

The radiation of the AGN population allows inference about the mass accretion onto their supermassive black holes. Marconi et al. (2004) showed that the SMBH relics found in the local Universe can be brought into agreement with the total luminosity output of AGN. In

their calculation, the Compton-thin luminosity function of Ueda et al. (2003) was extended by assuming a certain Compton-thick fraction. Without the latter, previous works on the luminosity function have found their low black hole mass densities underestimated due to low detection rates of Compton-thick AGN (e.g. Aird et al. 2010) or reported broad uncertainties in the fraction of Compton-thick AGN (e.g. Ueda et al. 2014). This work has placed tight constraints on both the number of obscured and Compton-thick AGN. It showed the accretion growth of SMBHs is indeed dominated by Compton-thick and Compton-thin AGN ($74^{+40}_{-5}\%$), while unobscured AGN only play a minor role. Compton-thick and Compton-thin AGN each contribute about half of the luminosity output (see Figure 7.3.2), with the Compton-thick AGN contributing $40^{+6}_{-6}\%$ of the total. This shows that Compton-thick AGN are an important contributor to the accretion history.

The overall evolution of the accretion density for the AGN population is presented in Figure 7.3.2. There, the peak is in a broad plateau at $z = 1 - 3$, with a drop-off towards high redshift. This trend may corroborate claims that the accretion onto SMBHs correlates with the star-formation history of galaxies (see Merloni & Heinz 2013). Previous works have claimed such a decline (Brusa et al. 2009; Civano et al. 2011; Lehmer et al. 2012; Vito et al. 2014; Ueda et al. 2014), however only the last work included the absorption of AGN luminosities. Also, previous works have relied on the LDDE parameterisation of the luminosity function evolution to determine the number density evolution (e.g. Ueda et al. 2014). In contrast, the results presented do not a-priori assume any form of the evolution at high redshift. Thus, this trend is robustly detected with the non-parametric methodology.

Figure 7.3.4 shows that the luminosity output due to accretion onto black holes are in agreement with observations of the black hole mass density in local AGN. This is only the case however if Compton-thin obscured AGN are included. Due to the large uncertainties, Compton-thick AGN are not required to explain the local black hole mass density, yet the quoted fraction of $\sim 35\%$ is compatible. Beyond Compton-thick AGN, another large fraction of additional “hidden” population, i.e. not X-ray detectable due to even higher obscuring columns, can also be excluded from the presented results. Such a population missed in X-ray surveys is suggested in some infrared surveys (see discussion in Chapter 6, in particular Section 6.3.1). Finally, recent updates on local scaling relationships by Kormendy & Ho (2013) indicate that the scaling normalisation of previous works may have been underestimated by a factor of a few, and that the local black hole mass density may need to be updated to a higher value. If this is the case, the Compton-thick AGN may indeed be required to explain the updated, local black hole mass density.

Some systematic uncertainties enter into the black hole mass density evolution, most notably that the black hole mass function and the spin distribution are poorly understood. The black hole mass and spin both influence the radiative efficiency of accretion onto the black hole. This work has assumed a fixed radiative efficiency of $\mu = 10\%$. For investigating the radiative efficiency, a number of additional physical effects have to be taken into account. The formalism of “thin disks” (Shakura & Sunyaev 1973) was extended to a modern treatment of “slim disks” (see Abramowicz & Fragile 2013, for a review), incorporating hydrodynamics, radiative feedback and magnetic fields. Sadowski (2011)

shows via simulations of relativistic accretion disks that beyond the Eddington accretion limit, the radiative efficiency decreases like a powerlaw with index $\gamma \approx -1$, due to photon trapping and quick heat flows into the black hole. Such a relation is illustrated in the bottom panel of Figure 7.3.1, for a mildly rotating black hole. The differences between assuming $\mu = 10\%$ and a more realistic, accretion-rate dependent conversion are shown in Figure 7.3.1. In general, the agreement is very good until the Eddington luminosity is reached. At very low accretion rates, different accretion modes are relevant, which also have low radiative efficiencies (see Section 3.2.1). It has also been suggested that also the bolometric corrections are dependent on the accretion rate, requiring a more complicated, accretion-state based treatment (Vasudevan & Fabian 2007).

The largest uncertainties however arise from the bright end of the luminosity function, where the data used in this work are sparse. The addition of a large survey is necessary to constrain the accretion history of the Universe better with the non-parametric technique. The best upcoming survey for this task is the all-sky survey performed by eROSITA (Predehl et al. 2014), expected to detect a few million, predominately bright, AGN. This is discussed further in Section 8.

7.5 Conclusions

This study has benefited from a non-parametric luminosity function analysis which make these conclusions particularly robust. Furthermore, this work has benefited from detailed study of the X-ray spectrum to yield robust intrinsic $2 - 10\text{keV}$ luminosities. The findings of this chapter with regard to the total number of AGN as a function of luminosity and redshift can be summarised as follows:

1. The luminosity function of AGN can be described as a bending powerlaw at all redshifts.
2. Only the normalisation and the luminosity break parameters evolve significantly. The increase in the latter indicates that massive black holes form earlier in the Universe, while at low redshift, predominantly small black holes accrete (anti-hierarchical growth, “downsizing”).
3. The luminosity output of the AGN population is dominated by a broad peak of AGN activity at $z = 0.5 - 3$. At lower redshifts, luminosity evolution reduces the luminosity output. Towards higher redshifts, the number density of AGN declines. Including obscured AGN, the total radiative output is sufficient to explain the black hole masses found in the local universe, completing the Soltan argument.

8 Summary

Active Galactic Nuclei are mysterious objects that live in the centres of massive galaxies. Some outshine their host galaxy with luminosities of billions of stars, which is more energy than that needed to unbind the galactic gravitational potential. It thus seems obvious that these objects can have significant effects on the formation and evolution of galaxies, and that the galaxy ecosystem can not be understood without them.

The first step for placing AGN in the context of galaxy formation is to count them, and to measure their total luminosity output. This enterprise is hampered by selection biases, not least due to the significant absorbing material seen in many AGN. Once the selection bias due to obscuration is understood, the fraction of obscured AGN can be determined. However, in this process it is essential that distances are known, by associating the X-ray sources with optical / infrared counterparts and obtaining spectra to measure the redshift of these sources. Additionally, this thesis develops methodological advances to account for errors in the possibly incomplete or ambiguous counterpart association process as well as those arising in the redshift determination. These uncertainties have been propagated into all other results, such as the luminosity function, which describes the total number and luminosity distribution of AGN over cosmic time.

The luminosity of AGN is thought to be powered by accretion onto super-massive black holes. One open question regarding the cosmic evolution of AGN has been when the large black hole masses found dormant in nearby galaxies have been assembled. Generally the Universe of eleven to five billion years ago is forming stars at record rates. As shown in this work and others, the peak of AGN activity appears also during this period, suggesting that a large amount of the gas not baked into new stars and planets fell into black holes. Already 3 billion years after the Big Bang, half of the final black hole masses have been accumulated. However, a large fraction of these AGN are hidden behind thick layers of gas and dust, a veil which this thesis lifts, in part thanks to advances in methodology. This thesis has provided new robust constraints using several, significant advances in methodology and by measuring the black hole growth also in the most heavily obscured, Compton-thick AGN. The luminosity function also provides important constraints to cosmological simulations, which try to explain the evolution of galaxies using recipes for the activation of SMBH accretion in the life cycle of galaxies.

Beyond population studies in the context of galaxies, AGN are interesting objects on their own right, because they allow glimpses on extreme environments around black holes, an exotic type of object. This thesis also studies the nature of the obscuring gas and dust through X-ray spectroscopy. Large amounts of internal reflection is seen, showing that the super-massive black hole is surrounded extensively by both thick and thin gas. This is consistent with the large fraction of obscured (75%) and heavily obscured (35%) AGN

detected. Furthermore, the obscuration appears to be dependent on the luminosity to black hole mass ratio. These findings have implications for distinguishing possible physical processes governing the formation and maintenance of these clouds.

Future work and observations

A limitation of the work presented here is the loose constraints on the space density of the most luminous AGN. These objects, although rare, nevertheless contribute substantially to the accretion density and therefore uncertainties in their number density have direct implication on our understanding of the accretion history of the Universe, the fraction of obscured AGN. Better constraints can be achieved either by using more assumptions (e.g. that the shape is a bending powerlaw) or by adding a wide-area survey (see below).

In the future I would like to investigate the history of obscured black hole growth in more detail, reconstructing not just the black hole mass density but also tracking the black hole mass function over cosmic time. This can be contrasted with attempts which use only type-1 AGN to reconstruct the black hole mass function at higher redshifts (e.g. Schulze et al. 2015). With the fueling function (at each redshift, the amount of mass accreted at each mass of the black hole mass function) constrained, it is possible to predict the star formation of galaxies based on the shared-reservoir hypothesis (see e.g. Hickox et al. 2014), and compare with the observed star formation history. Furthermore, recipes for the creation of the obscurer as a function of Eddington rate can be tested (e.g. Wada 2012), and compared to the column density distributions obtained. This will constrain physical models for the maintenance of the torus and determine whether galactic absorption is important. Finally, comparing the luminosity function and column density distribution to XRB measurements would be important to settle the long-standing disagreements e.g. in the Compton-thick fraction. For this comparison, the distribution of the relative normalisation of additional spectral components, i.e. Compton reflection and soft scattering powerlaw, should be re-used. Furthermore, simulating a torus with a column density gradient would be useful to obtain a self-consistent modelling of the Compton scattering.

The constraints on e.g. the black hole mass density are relatively broad due to poor constraints on the number of bright AGN. To account for these rare AGN, the analysis would benefit from the addition of a large, shallow field. The eROSITA mission (Predehl et al. 2014), scheduled to launch in 2016, will perform an all-sky survey over four years (eRASS). I computed the sensitivity curve of this survey to point sources, taking into account the expected instrument response, proposed mapping strategy and expected background (Figure 8.0.1, published in Merloni et al. 2012). The flux sensitivity in the hard band is comparable to that adopted for the XMM-XXL survey, but the survey is 2000 times larger. Based on the non-parametric luminosity function, a total number of AGN detected is 3.9 million, including 150,000-250,000 Compton-thick AGN in the redshift range $z = 0.5 - 3$. The eRASS AGN sample will be invaluable to study obscured AGN and their host galaxies in detail. Figure 8.0.2 shows the predicted number of sources as a function of their flux. Because of the large uncertainties at the bright end, the luminosity function

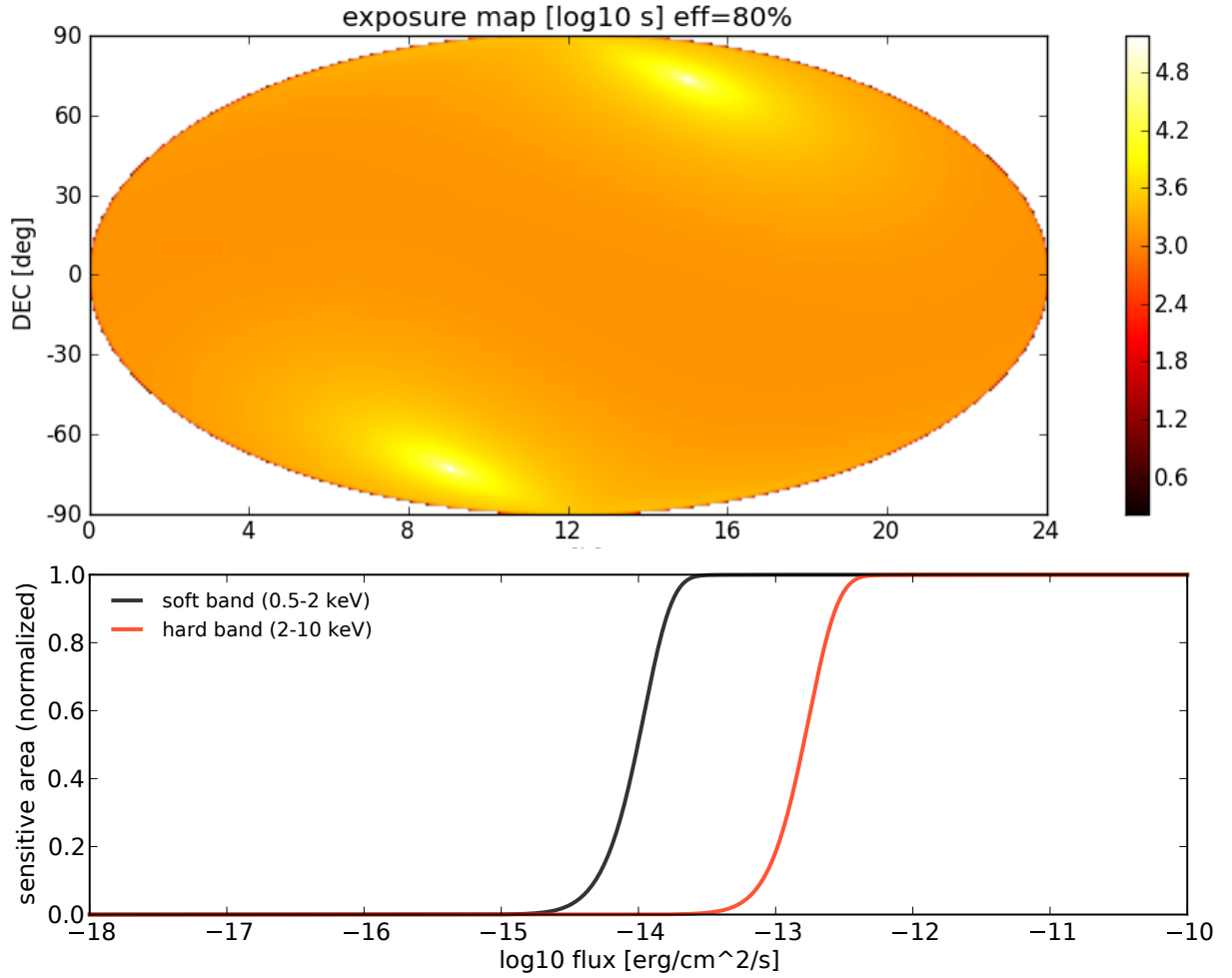


Figure 8.0.1: Top panel: The exposure map shows that the mapping strategy gives longer exposure to the ecliptic poles. Bottom panel: Sensitivity curves for the eROSITA survey. The survey will be most sensitive in the soft band (black line).

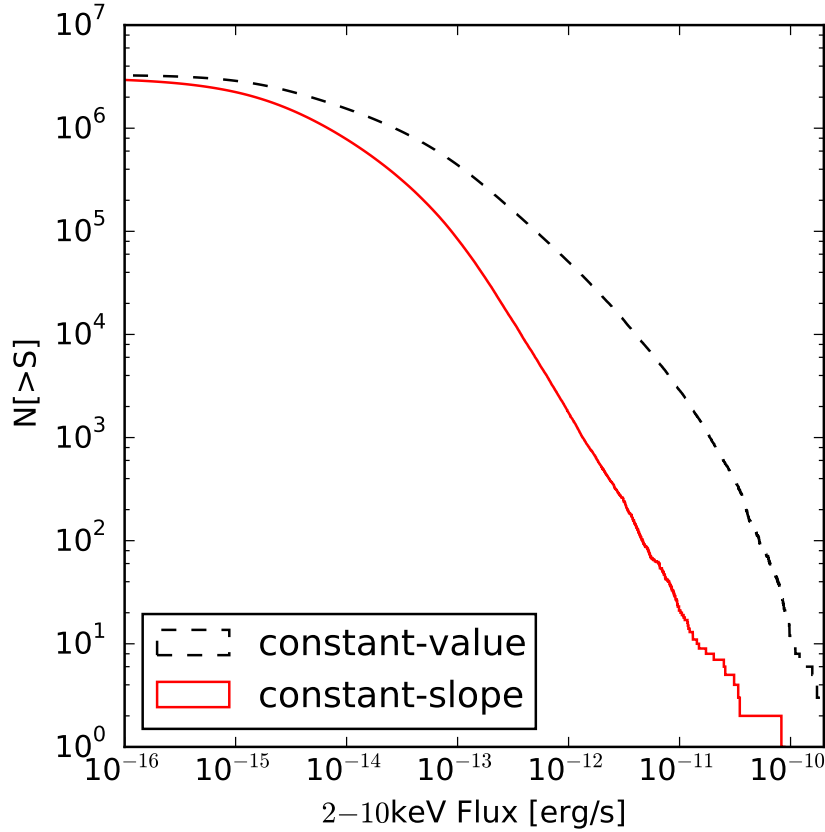


Figure 8.0.2: Cumulative number of detected sources predicted for the eROSITA all-sky survey as a function of flux in the observed 2 – 10keV band. The number of sources above a certain flux are plotted. The non-parametric luminosity function based on the constant-value prior (black dashed line) predicts a large number of very bright AGN, while the results based on the constant-slope prior (red solid line) predicts a large number of faint AGN. This shows that the eROSITA survey is highly sensitive to the bright-end, and will place constraints on these the uncertainties left open in this thesis even before redshift follow-up is complete.

determined in this work predicts a wide range (two orders of magnitude) for the number of bright AGN for eROSITA. In other words, eROSITA is highly sensitive to determine the bright end by the number of detections alone, even without counter-part followup.

Another limitation of this work is the small count statistics of X-ray spectra of high-redshift sources which limits the analysis of their obscuration. The ATHENA+ mission (Nandra et al. 2013) is a planned next-generation telescope on the road to replace the Chandra and XMM-Newton telescopes in 2028. Some technical highlights of this mission include a substantially larger sensitive area and more sensitive detectors. With this instrument it will be possible to study, for example, the nature of the soft scattering component. In this work it has been argued that ionised clouds reflect the intrinsic radiation past the obscurer. The absorption lines of these ionised clouds would become observable with the high-resolution spectrographs on-board ATHENA+. Also, if these clouds are part of bulk outflows, the blueshift of these can be measured and will give detailed information about the energy deposited into the host galaxies. Furthermore it will be interesting to study the obscurer in more detail. For instance, the metallicity of high-redshift obscurers is poorly understood, and remains a systematic uncertainty in the geometry of the obscurer, but could potentially be constrained by detailed observations.

I have demonstrated a technique to estimate redshifts for obscured AGN based on their X-ray spectrum alone. However, for eROSITA, the fraction of sources in which this technique will succeed is low, due to the preferential selection of bright, unobscured AGN. Large optical / IR follow-up campaigns are thus essential (ground and space-based). Also for the ATHENA+ observations of deep fields it is essential to find counterparts by high-resolution, deep exposures and to determine their redshifts (e.g. James Webb Space Telescope).

9 Acknowledgements

The doctoral graduation finalises my long journey as a student, and I would like to thank the many people who have encouraged and supported me. Foremost, I am most grateful to my parents. I would like to thank Dr. Katrin Seyr, Prof. Werner W. Weiss (who introduced me to research in astronomy), Dr. Michael Gruberbauer (who introduced me to Bayesian methods), Prof. Sergei Gulyaev (for his kind welcoming to his institute), Tim Natusch (for his patience in observing with me in the land of the long white cloud and always being available for answering questions – please take some holidays!).

My gratitude goes to Dr. Antonis Georgakakis, who knows every paper in existence, and who shared his never-ending patience, knowledge and kindness in my doctoral day-to-day supervision – you are an amazing supervisor! I am thankful to Prof. Paul Nandra who wisely chose and tailored my research area to my interests so that in the end I got good results, Dr. Andrea Merloni (for extended discussions) and Dr. Mara Salvato (for her enthusiasm for science and her kindness). I have always experienced our research group as simultaneously humane and professional, which is the best combination.

I thank my collaborators, specifically for the painstaking work of optical identification and redshift estimation performed by Li-Ting Hsu, Marie-Luise Menzel, Dr. Mara Salvato and Dr. James Aird, as well as Dr. Murray Brightman and Dr. Cyprian Rangel for X-ray detection and extraction of X-ray spectra. I would also like to thank Dr. Mike Anderson for proof-reading this thesis (remaining errors are mine).

The environment in Garching has been extremely helpful in broadening my horizon for astronomical research through a myriad of talks and discussions with too many people to name. Thank you!

Bibliography

- Abramowicz, M. A. & Fragile, P. C. 2013, *Living Reviews in Relativity*, 16, 1
- Aihara, H., Allende Prieto, C., An, D., et al. 2011, *ApJS*, 193, 29
- Aird, J., Nandra, K., Georgakakis, A., et al. 2008, *MNRAS*, 387, 883
- Aird, J., Nandra, K., Laird, E. S., et al. 2010, *MNRAS*, 401, 2531
- Akaike, H. 1974, *Automatic Control, IEEE Transactions on*, 19, 716
- Akylas, A., Georgakakis, A., Georgantopoulos, I., Brightman, M., & Nandra, K. 2012, *A&A*, 546, A98
- Akylas, A. & Georgantopoulos, I. 2008, *A&A*, 479, 735
- Akylas, A. & Georgantopoulos, I. 2009, *A&A*, 500, 999
- Akylas, A., Georgantopoulos, I., Georgakakis, A., Kitsionas, S., & Hatziminaoglou, E. 2006, *A&A*, 459, 693
- Alexander, D. M., Bauer, F. E., Brandt, W. N., et al. 2011, *ApJ*, 738, 44
- Alexander, D. M., Stern, D., Del Moro, A., et al. 2013, *ApJ*, 773, 125
- Allevato, V., Finoguenov, A., Civano, F., et al. 2014, *ApJ*, 796, 4
- Antonucci, R. 1993, *ARA&A*, 31, 473
- Antonucci, R. R. J. 1982, *Nature*, 299, 605
- Antonucci, R. R. J. & Miller, J. S. 1985, *ApJ*, 297, 621
- Arévalo, P., Bauer, F. E., Puccetti, S., et al. 2014, *ApJ*, 791, 81
- Assef, R. J., Stern, D., Kochanek, C. S., et al. 2013, *ApJ*, 772, 26
- Awaki, H., Koyama, K., Inoue, H., & Halpern, J. P. 1991, *PASJ*, 43, 195
- Baade, W. & Minkowski, R. 1954, *ApJ*, 119, 215

- Baganoff, F. 1999, ACIS On-orbit Background Rates and Spectra from Chandra OAC Phase 1, ACIS Memo 162, Massachusetts Institute of Technology, Center for Space Research
- Baldwin, J. A., Phillips, M. M., & Terlevich, R. 1981, *PASP*, 93, 5
- Barger, A. J. & Cowie, L. L. 2005, *ApJ*, 635, 115
- Barro, G., Pérez-González, P. G., Gallego, J., et al. 2011a, *ApJS*, 193, 13
- Barro, G., Pérez-González, P. G., Gallego, J., et al. 2011b, *ApJS*, 193, 30
- Bassani, L., Dadina, M., Maiolino, R., et al. 1999, *ApJS*, 121, 473
- Bassani, L., Molina, M., Malizia, A., et al. 2006, *ApJ*, 636, L65
- Bauer, F. E., Arevalo, P., Walton, D. J., et al. 2014, *ArXiv e-prints*
- Bayarri, M. J. & Castellanos, M. E. 2008, *ArXiv e-prints*
- Beasley, A. J., Murowinski, R., & Tarenghi, M. 2006, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Vol. 6267, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 2
- Bechtold, J., Czerny, B., Elvis, M., Fabbiano, G., & Green, R. F. 1987, *ApJ*, 314, 699
- Blandford, R. D. & Begelman, M. C. 1999, *MNRAS*, 303, L1
- Bolton, A. S., Schlegel, D. J., Aubourg, É., et al. 2012, *AJ*, 144, 144
- Bongiorno, A., Mignoli, M., Zamorani, G., et al. 2010, *A&A*, 510, A56
- Bournaud, F., Dekel, A., Teyssier, R., et al. 2011, *ApJ*, 741, L33
- Bournaud, F., Elmegreen, B. G., & Elmegreen, D. M. 2007, *ApJ*, 670, 237
- Boyle, B. J., Griffiths, R. E., Shanks, T., Stewart, G. C., & Georgantopoulos, I. 1993, *MNRAS*, 260, 49
- Boyle, B. J. & Terlevich, R. J. 1998, *MNRAS*, 293, L49
- Brandt, W. N. & Hasinger, G. 2005, *ARA&A*, 43, 827
- Brightman, M. & Nandra, K. 2011a, *MNRAS*, 413, 1206
- Brightman, M. & Nandra, K. 2011b, *MNRAS*, 414, 3084
- Brightman, M. & Nandra, K. 2012, *MNRAS*, 422, 1166
- Brightman, M., Nandra, K., Salvato, M., et al. 2014, *MNRAS*, 443, 1999

- Brightman, M. & Ueda, Y. 2012, MNRAS, 423, 702
- Broos, P. S., Townsley, L. K., Feigelson, E. D., et al. 2010, ApJ, 714, 1582
- Brusa, M., Comastri, A., Daddi, E., et al. 2005, A&A, 432, 69
- Brusa, M., Comastri, A., Gilli, R., et al. 2009, ApJ, 693, 8
- Brusa, M., Zamorani, G., Comastri, A., et al. 2007, ApJS, 172, 353
- Buchner, J., Georgakakis, A., Nandra, K., et al. 2015, ApJ, 802, 89
- Buchner, J., Georgakakis, A., Nandra, K., et al. 2014, A&A, 564, A125
- Budavári, T. & Szalay, A. S. 2008, ApJ, 679, 301
- Burlon, D., Ajello, M., Greiner, J., et al. 2011, ApJ, 728, 58
- Burtscher, L., Meisenheimer, K., Tristram, K. R. W., et al. 2013, A&A, 558, A149
- Capak, P., Aussel, H., Ajiki, M., et al. 2007, ApJS, 172, 99
- Cappelluti, N., Brusa, M., Hasinger, G., et al. 2009, A&A, 497, 635
- Cappi, M., Panessa, F., Bassani, L., et al. 2006, A&A, 446, 459
- Cardamone, C. N., van Dokkum, P. G., Urry, C. M., et al. 2010, ApJS, 189, 270
- Cash, W. 1979, ApJ, 228, 939
- Cattaneo, A., Faber, S. M., Binney, J., et al. 2009, Nature, 460, 213
- Ciotti, L. & Ostriker, J. P. 1997, ApJ, 487, L105
- Ciotti, L. & Ostriker, J. P. 2001, ApJ, 551, 131
- Civano, F., Brusa, M., Comastri, A., et al. 2011, ApJ, 741, 91
- Civano, F., Elvis, M., Brusa, M., et al. 2012, ApJS, 201, 30
- Coil, A. L., Georgakakis, A., Newman, J. A., et al. 2009, ApJ, 701, 1484
- Comastri, A., Ranalli, P., Iwasawa, K., et al. 2011, A&A, 526, L9
- Cooper, M. C., Aird, J. A., Coil, A. L., et al. 2011, ApJS, 193, 14
- Cooper, M. C., Griffith, R. L., Newman, J. A., et al. 2012, MNRAS, 419, 3018
- Cowie, L. L., Songaila, A., Hu, E. M., & Cohen, J. G. 1996, AJ, 112, 839
- Croom, S. M., Richards, G. T., Shanks, T., et al. 2009, MNRAS, 399, 1755

- Croton, D. J., Springel, V., White, S. D. M., et al. 2006, MNRAS, 365, 11
- Czerny, B. & Elvis, M. 1987, ApJ, 321, 305
- Davis, M., Guhathakurta, P., Konidaris, N. P., et al. 2007, ApJ, 660, L1
- Dawson, K. S., Schlegel, D. J., Ahn, C. P., et al. 2013, AJ, 145, 10
- de Rosa, A., Panessa, F., Bassani, L., et al. 2012, MNRAS, 420, 2087
- Dickey, J. M. & Lockman, F. J. 1990, ARA&A, 28, 215
- Doeleman, S. S., Fish, V. L., Schenck, D. E., et al. 2012, Science, 338, 355
- Donley, J. L., Koekemoer, A. M., Brusa, M., et al. 2012, ApJ, 748, 142
- Donley, J. L., Rieke, G. H., Alexander, D. M., Egami, E., & Pérez-González, P. G. 2010, ApJ, 719, 1393
- Ebrero, J., Carrera, F. J., Page, M. J., et al. 2009, A&A, 493, 55
- Eisenstein, D. J., Weinberg, D. H., Agol, E., et al. 2011, AJ, 142, 72
- Elitzur, M. 2008, Mem. Soc. Astron. Italiana, 79, 1124
- Elitzur, M. & Ho, L. C. 2009, ApJ, 701, L91
- Elitzur, M. & Shlosman, I. 2006, ApJ, 648, L101
- Elvis, M., Civano, F., Vignali, C., et al. 2009, ApJS, 184, 158
- Elvis, M., Wilkes, B. J., McDowell, J. C., et al. 1994, ApJS, 95, 1
- Enoki, M., Ishiyama, T., Kobayashi, M. A. R., & Nagashima, M. 2014, ApJ, 794, 69
- Fabian, A. C. 1989, in ESA Special Publication, Vol. 296, Two Topics in X-Ray Astronomy, Volume 1: X Ray Binaries. Volume 2: AGN and the X Ray Background, ed. J. Hunt & B. Battrick, 1097–1104
- Fabian, A. C. 1999, MNRAS, 308, L39
- Fabian, A. C. & Barcons, X. 1992, ARA&A, 30, 429
- Fanidakis, N., Baugh, C. M., Benson, A. J., et al. 2012, MNRAS, 419, 2797
- Feroz, F., Hobson, M. P., & Bridges, M. 2009, MNRAS, 398, 1601
- Ferrarese, L. & Ford, H. 2005, Space Sci. Rev., 116, 523
- Ferrarese, L., Ford, H. C., & Jaffe, W. 1996, ApJ, 470, 444

- Ferrarese, L. & Merritt, D. 2000, *ApJ*, 539, L9
- Fiore, F., Grazian, A., Santini, P., et al. 2008, *ApJ*, 672, 94
- Fiore, F., Puccetti, S., Brusa, M., et al. 2009, *ApJ*, 693, 447
- Fontanot, F., De Lucia, G., Monaco, P., Somerville, R. S., & Santini, P. 2009, *MNRAS*, 397, 1776
- Freeman, P., Doe, S., & Siemiginowska, A. 2001, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Vol. 4477, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, ed. J.-L. Starck & F. D. Murtagh, 76–87
- Gabor, J. M., Impey, C. D., Jahnke, K., et al. 2009, *ApJ*, 691, 705
- Garmire, G. P., Bautz, M. W., Ford, P. G., Nousek, J. A., & Ricker, Jr., G. R. 2003, *Advanced CCD imaging spectrometer (ACIS) instrument on the Chandra X-ray Observatory*
- Gaskell, C. M., Goosmann, R. W., & Klimek, E. S. 2008, *Mem. Soc. Astron. Italiana*, 79, 1090
- Gebhardt, K., Bender, R., Bower, G., et al. 2000, *ApJ*, 539, L13
- Genzel, R., Eisenhauer, F., & Gillessen, S. 2010, *Reviews of Modern Physics*, 82, 3121
- Georgakakis, A., Mountrichas, G., Salvato, M., et al. 2014, *MNRAS*, 443, 3327
- Georgakakis, A. & Nandra, K. 2011, *MNRAS*, 414, 992
- Georgakakis, A., Nandra, K., Laird, E. S., Aird, J., & Trichas, M. 2008, *MNRAS*, 388, 1205
- Georgakakis, A., Rowan-Robinson, M., Nandra, K., et al. 2010, *MNRAS*, 406, 420
- Georgantopoulos, I., Comastri, A., Vignali, C., et al. 2013, *ArXiv e-prints*
- George, I. M. & Fabian, A. C. 1991, *MNRAS*, 249, 352
- Ghisellini, G., Haardt, F., & Fabian, A. C. 1993, *MNRAS*, 263, L9
- Giacconi, R., Gursky, H., Paolini, F. R., & Rossi, B. B. 1962, *Physical Review Letters*, 9, 439
- Gierliński, M. & Done, C. 2004, *MNRAS*, 349, L7
- Gilli, R., Comastri, A., & Hasinger, G. 2007, *A&A*, 463, 79
- Gilli, R., Vignali, C., Mignoli, M., et al. 2010, *A&A*, 519, A92

- Gondhalekar, P. M., Rouillon-Foley, C., & Kellett, B. J. 1997, *MNRAS*, 288, 260
- Grogin, N. A., Conselice, C. J., Chatzichristou, E., et al. 2005, *ApJ*, 627, L97
- Grogin, N. A., Koekemoer, A. M., Schreier, E. J., et al. 2003, *ApJ*, 595, 685
- Guainazzi, M., Matt, G., & Perola, G. C. 2005, *A&A*, 444, 119
- Gunn, J. E., Siegmund, W. A., Mannery, E. J., et al. 2006, *AJ*, 131, 2332
- Guo, Y., Ferguson, H. C., Giavalisco, M., et al. 2013, *ApJS*, 207, 24
- Häring, N. & Rix, H.-W. 2004, *ApJ*, 604, L89
- Hasinger, G. 2008, *A&A*, 490, 905
- Hasinger, G., Miyaji, T., & Schmidt, M. 2005, *A&A*, 441, 417
- Heckman, T. M. 1980, *A&A*, 87, 152
- Hickox, R. C., Mullaney, J. R., Alexander, D. M., et al. 2014, *ApJ*, 782, 9
- Hirschmann, M., Dolag, K., Saro, A., et al. 2014, *MNRAS*, 442, 2304
- Hirschmann, M., Naab, T., Somerville, R. S., Burkert, A., & Oser, L. 2012, *MNRAS*, 419, 3200
- Ho, L. C., Filippenko, A. V., & Sargent, W. L. W. 1997, in *Astronomical Society of the Pacific Conference Series*, Vol. 113, IAU Colloq. 159: Emission Lines in Active Galaxies: New Methods and Techniques, ed. B. M. Peterson, F.-Z. Cheng, & A. S. Wilson, 429
- Hoaglin, D. C., Mosteller, F., & Tukey, J. W. 1983, *Understanding robust and exploratory data analysis* (Wiley New York)
- Hoffman, M. D. & Gelman, A. 2011, *ArXiv e-prints*
- Hönig, S. F. & Beckert, T. 2007, *MNRAS*, 380, 1172
- Hopkins, P. F., Hayward, C. C., Narayanan, D., & Hernquist, L. 2012, *MNRAS*, 420, 320
- Hopkins, P. F., Hernquist, L., Cox, T. J., et al. 2005a, *ApJ*, 630, 705
- Hopkins, P. F., Hernquist, L., Cox, T. J., et al. 2005b, *ApJ*, 630, 716
- Hopkins, P. F., Hernquist, L., Cox, T. J., et al. 2006a, *ApJS*, 163, 1
- Hopkins, P. F., Hernquist, L., Cox, T. J., et al. 2006b, *ApJ*, 639, 700
- Hopkins, P. F., Richards, G. T., & Hernquist, L. 2007, *ApJ*, 654, 731

- Hopkins, P. F., Somerville, R. S., Hernquist, L., et al. 2006c, *ApJ*, 652, 864
- Hsieh, B.-C., Wang, W.-H., Hsieh, C.-C., et al. 2012, *ApJS*, 203, 23
- Hsu, L.-T., Salvato, M., Nandra, K., et al. 2014, *ApJ*, 796, 60
- Hunter, J. D. 2007, *Computing In Science & Engineering*, 9, 90
- Ilbert, O., Arnouts, S., McCracken, H. J., et al. 2006, *A&A*, 457, 841
- Ilbert, O., Capak, P., Salvato, M., et al. 2009, *ApJ*, 690, 1236
- Iwasawa, K., Gilli, R., Vignali, C., et al. 2012, *A&A*, 546, A84
- Jansen, F., Lumb, D., Altieri, B., et al. 2001, *A&A*, 365, L1
- Jeffreys, H. 1998, *International series of monographs on physics*.
- Jia, J., Ptak, A., Heckman, T., & Zakamska, N. L. 2013, *ApJ*, 777, 27
- Jones, E., Oliphant, T., Peterson, P., et al. 2001–, *SciPy: Open source scientific tools for Python*
- Kalberla, P. M. W., Burton, W. B., Hartmann, D., et al. 2005, *A&A*, 440, 775
- Kalfountzou, E., Civano, F., Elvis, M., Trichas, M., & Green, P. 2014, *MNRAS*, 445, 1430
- Katz, J. I. 1976, *ApJ*, 206, 910
- Kauffmann, G., Heckman, T. M., Tremonti, C., et al. 2003, *MNRAS*, 346, 1055
- Kelly, B. C., Fan, X., & Vestergaard, M. 2008, *ApJ*, 682, 874
- Khachikian, E. Y. & Weedman, D. W. 1974, *ApJ*, 192, 581
- King, A. 2003, *ApJ*, 596, L27
- King, A. R. & Pringle, J. E. 2006, *MNRAS*, 373, L90
- King, A. R., Pringle, J. E., & Hofmann, J. A. 2008, *MNRAS*, 385, 1621
- Kocevski, D. D., Faber, S. M., Mozena, M., et al. 2011, *ArXiv e-prints*
- Kormendy, J. & Ho, L. C. 2013, *ARA&A*, 51, 511
- Krolik, J. H. & Begelman, M. C. 1988, *ApJ*, 329, 702
- Krolik, J. H. & Kriss, G. A. 1995, *ApJ*, 447, 512
- Kullback, S. & Leibler, R. A. 1951, *The Annals of Mathematical Statistics*, 22, 79

- La Franca, F., Fiore, F., Comastri, A., et al. 2005, *ApJ*, 635, 864
- Lacy, M., Storrie-Lombardi, L. J., Sajina, A., et al. 2004, *ApJS*, 154, 166
- Laird, E. S., Nandra, K., Georgakakis, A., et al. 2009, *ApJS*, 180, 102
- Lanzuisi, G., Ranalli, P., Georgantopoulos, I., et al. 2014, *ArXiv e-prints*
- Lawrence, A. 1991, *MNRAS*, 252, 586
- Lawrence, A. & Elvis, M. 1982, *ApJ*, 256, 410
- Lehmer, B. D., Xue, Y. Q., Brandt, W. N., et al. 2012, *ApJ*, 752, 46
- Lilly, S. J., Le Brun, V., Maier, C., et al. 2009, *ApJS*, 184, 218
- Lored, T. J. 2004, in *American Institute of Physics Conference Series*, Vol. 735, American Institute of Physics Conference Series, ed. R. Fischer, R. Preuss, & U. V. Toussaint, 195–206
- Lusso, E., Comastri, A., Simmons, B. D., et al. 2012, *MNRAS*, 425, 623
- Lusso, E., Hennawi, J. F., Comastri, A., et al. 2013, *ApJ*, 777, 86
- Lynden-Bell, D. 1969, *Nature*, 223, 690
- Maccacaro, T., della Ceca, R., Gioia, I. M., et al. 1991, *ApJ*, 374, 117
- Madau, P. & Dickinson, M. 2014, *ARA&A*, 52, 415
- Magdziarz, P. & Zdziarski, A. A. 1995, *MNRAS*, 273, 837
- Magorrian, J., Tremaine, S., Richstone, D., et al. 1998, *AJ*, 115, 2285
- Maiolino, R. & Rieke, G. H. 1995, *ApJ*, 454, 95
- Malizia, A., Bassani, L., Bazzano, A., et al. 2012, *MNRAS*, 426, 1750
- Marconi, A., Risaliti, G., Gilli, R., et al. 2004, *MNRAS*, 351, 169
- Markowitz, A. G., Krumpe, M., & Nikutta, R. 2014, *MNRAS*, 439, 1403
- Marshall, H. L., Tananbaum, H., Avni, Y., & Zamorani, G. 1983, *ApJ*, 269, 35
- Mathez, G. 1976, *A&A*, 53, 15
- Matt, G. 2000, *A&A*, 355, L31
- Matt, G. 2002, *MNRAS*, 337, 147
- Matt, G., Brandt, W. N., & Fabian, A. C. 1996, *MNRAS*, 280, 823

- Matt, G., Fabian, A. C., Guainazzi, M., et al. 2000, MNRAS, 318, 173
- Matt, G., Fabian, A. C., & Reynolds, C. S. 1997, MNRAS, 289, 175
- Mayo, J. H. & Lawrence, A. 2013, MNRAS, 434, 1593
- McCracken, H. J., Capak, P., Salvato, M., et al. 2010, ApJ, 708, 202
- McHardy, I. M. 1989, in ESA Special Publication, Vol. 296, Two Topics in X-Ray Astronomy, Volume 1: X Ray Binaries. Volume 2: AGN and the X Ray Background, ed. J. Hunt & B. Battrick, 1111–1124
- Menci, N., Fiore, F., Puccetti, S., & Cavaliere, A. 2008, ApJ, 686, 219
- Merloni, A. 2004, MNRAS, 353, 1035
- Merloni, A., Bongiorno, A., Brusa, M., et al. 2014, MNRAS, 437, 3550
- Merloni, A. & Heinz, S. 2008, MNRAS, 388, 1011
- Merloni, A. & Heinz, S. 2013, Evolution of Active Galactic Nuclei, ed. T. D. Oswalt & W. C. Keel (Springer Netherlands), 503–566
- Merloni, A., Predehl, P., Becker, W., et al. 2012, ArXiv e-prints
- Mignoli, M., Vignali, C., Gilli, R., et al. 2013, A&A, 556, A29
- Miyaji, T., Hasinger, G., & Schmidt, M. 2000, A&A, 353, 25
- Molinari, S., Bally, J., Noriega-Crespo, A., et al. 2011, ApJ, 735, L33
- Morrison, R. & McCammon, D. 1983, ApJ, 270, 119
- Mullaney, J. R., Daddi, E., Béthermin, M., et al. 2012, ApJ, 753, L30
- Murphy, K. D. & Yaqoob, T. 2009, MNRAS, 397, 1549
- Mushotzky, R. 1984, in NASA Conference Publication, Vol. 2349, NASA Conference Publication, ed. J. M. Mead, R. D. Chapman, & Y. Kondo, 42–50
- Mushotzky, R. F. 1982, ApJ, 256, 92
- Nandra, K., Barret, D., Barcons, X., et al. 2013, ArXiv e-prints
- Nandra, K. & George, I. M. 1994, MNRAS, 267, 974
- Nandra, K., Laird, E. S., Adelberger, K., et al. 2005, MNRAS, 356, 568
- Nandra, K., O’Neill, P. M., George, I. M., & Reeves, J. N. 2007a, MNRAS, 382, 194

- Nandra, K., O'Neill, P. M., George, I. M., & Reeves, J. N. 2007b, *MNRAS*, 382, 194
- Nandra, K. & Pounds, K. A. 1994, *MNRAS*, 268, 405
- Narayan, R. & Yi, I. 1994, *ApJ*, 428, L13
- Narayan, R. & Yi, I. 1995a, *ApJ*, 444, 231
- Narayan, R. & Yi, I. 1995b, *ApJ*, 452, 710
- Nenkova, M., Sirocky, M. M., Nikutta, R., Ivezić, Ž., & Elitzur, M. 2008, *ApJ*, 685, 160
- Newman, A. B., Ellis, R. S., Bundy, K., & Treu, T. 2012, *ApJ*, 746, 162
- Norman, C., Hasinger, G., Giacconi, R., et al. 2002, *ApJ*, 571, 218
- Norman, C., Ptak, A., Hornschemeier, A., et al. 2004, *ApJ*, 607, 721
- Ogle, P. M., Brookings, T., Canizares, C. R., Lee, J. C., & Marshall, H. L. 2003, *A&A*, 402, 849
- Osmer, P. S. 1982, *ApJ*, 253, 28
- Panessa, F., Bassani, L., Cappi, M., et al. 2006, *A&A*, 455, 173
- Pérez-González, P. G., Rieke, G. H., Villar, V., et al. 2008, *ApJ*, 675, 234
- Perola, G. C., Matt, G., Cappi, M., et al. 2002, *A&A*, 389, 802
- Pier, E. A. & Krolik, J. H. 1992, *ApJ*, 399, L23
- Pierce, C. M., Lotz, J. M., Laird, E. S., et al. 2007, *ApJ*, 660, L19
- Predehl, P., Andritschke, R., Becker, W., et al. 2014, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Vol. 9144, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, 1
- Prieto, M. A., Mezcua, M., Fernández-Ontiveros, J. A., & Schartmann, M. 2014, *MNRAS*, 442, 2145
- Pringle, J. E. & Rees, M. J. 1972, *A&A*, 21, 1
- Protassov, R., van Dyk, D. A., Connors, A., Kashyap, V. L., & Siemiginowska, A. 2002, *ApJ*, 571, 545
- Ptak, A., Serlemitsos, P., Yaqoob, T., & Mushotzky, R. 1999, *ApJS*, 120, 179
- Rangel, C., Nandra, K., Laird, E. S., & Orange, P. 2013a, *MNRAS*
- Rangel, C., Nandra, K., Laird, E. S., & Orange, P. 2013b, *MNRAS*, 428, 3089

- Reichert, G. A., Mushotzky, R. F., Holt, S. S., & Petre, R. 1985, *ApJ*, 296, 69
- Reynolds, C. S. & Fabian, A. C. 1995, *MNRAS*, 273, 1167
- Richards, G. T., Lacy, M., Storrie-Lombardi, L. J., et al. 2006a, *ApJS*, 166, 470
- Richards, G. T., Strauss, M. A., Fan, X., et al. 2006b, *AJ*, 131, 2766
- Richstone, D., Ajhar, E. A., Bender, R., et al. 1998, *Nature*, 395, A14
- Risaliti, G., Elvis, M., Fabbiano, G., et al. 2007, *ApJ*, 659, L111
- Risaliti, G., Elvis, M., & Nicastro, F. 2002, *ApJ*, 571, 234
- Risaliti, G., Maiolino, R., & Salvati, M. 1999, *ApJ*, 522, 157
- Rivers, E., Markowitz, A., & Rothschild, R. 2013, *ApJ*, 772, 114
- Rosario, D. J., Mozena, M., Wuyts, S., et al. 2011, *ArXiv e-prints*
- Rosario, D. J., Santini, P., Lutz, D., et al. 2013, *ApJ*, 771, 63
- Roseboom, I. G., Lawrence, A., Elvis, M., et al. 2013, *MNRAS*, 429, 1494
- Ross, R. R. & Fabian, A. C. 1993, *MNRAS*, 261, 74
- Sadowski, A. 2011, *ArXiv e-prints*
- Saintonge, A., Kauffmann, G., Kramer, C., et al. 2011a, *MNRAS*, 415, 32
- Saintonge, A., Kauffmann, G., Wang, J., et al. 2011b, *MNRAS*, 415, 61
- Salpeter, E. E. 1964, *ApJ*, 140, 796
- Salvato, M., Hasinger, G., Ilbert, O., et al. 2009, *ApJ*, 690, 1250
- Salvato, M., Ilbert, O., Hasinger, G., et al. 2011, *ApJ*, 742, 61
- Sanders, D. B., Salvato, M., Aussel, H., et al. 2007, *ApJS*, 172, 86
- Sanders, D. B., Soifer, B. T., Elias, J. H., et al. 1988, *ApJ*, 325, 74
- Santini, P., Rosario, D. J., Shao, L., et al. 2012, *A&A*, 540, A109
- Sarzi, M., Shields, J. C., Schawinski, K., et al. 2010, *MNRAS*, 402, 2187
- Sazonov, S., Revnivtsev, M., Burenin, R., et al. 2008, *A&A*, 487, 509
- Schmidt, M. 1968, *ApJ*, 151, 393
- Schmidt, M. & Green, R. F. 1983, *ApJ*, 269, 352

- Schmidt, M., Schneider, D. P., & Gunn, J. E. 1995, *AJ*, 110, 68
- Schulze, A., Bongiorno, A., Gavignaud, I., et al. 2015, *MNRAS*, 447, 2085
- Schwarz, G. 1978, *Ann. Stat.*, 6, 461
- Scoville, N., Aussel, H., Brusa, M., et al. 2007, *ApJS*, 172, 1
- Setti, G. & Woltjer, L. 1989, *A&A*, 224, L21
- Seyfert, C. K. 1943, *ApJ*, 97, 28
- Shakura, N. I. & Sunyaev, R. A. 1973, *A&A*, 24, 337
- Shankar, F., Salucci, P., Granato, G. L., De Zotti, G., & Danese, L. 2004, *MNRAS*, 354, 1020
- Shi, Y., Helou, G., & Armus, L. 2013, *ApJ*, 777, 6
- Silk, J. & Rees, M. J. 1998, *A&A*, 331, L1
- Silverman, J. D., Green, P. J., Barkhouse, W. A., et al. 2008, *ApJ*, 679, 118
- Simpson, C. 2005, *MNRAS*, 360, 565
- Sinharay, S. & Stern, H. S. 2003, *Journal of Statistical Planning and Inference*, 111, 209, special issue I: Model Selection, Model Diagnostics, Empirical Bayes and Hierarchical Bayes
- Skilling, J. 2004, in *AIP Conference Proceedings*, Vol. 735, 395
- Smee, S. A., Gunn, J. E., Uomoto, A., et al. 2013, *AJ*, 146, 32
- Soltan, A. 1982, *MNRAS*, 200, 115
- Somerville, R. S., Hopkins, P. F., Cox, T. J., Robertson, B. E., & Hernquist, L. 2008, *MNRAS*, 391, 481
- Stalin, C. S., Petitjean, P., Srianand, R., et al. 2010, *MNRAS*, 401, 294
- Stan Development Team. 2014, *Stan: A C++ Library for Probability and Sampling*, Version 2.2
- Stark, A. A., Gammie, C. F., Wilson, R. W., et al. 1992, *ApJS*, 79, 77
- Steffen, A. T., Barger, A. J., Cowie, L. L., Mushotzky, R. F., & Yang, Y. 2003, *ApJ*, 596, L23
- Stern, D., Assef, R. J., Benford, D. J., et al. 2012, *ApJ*, 753, 30

- Strüder, L., Briel, U., Dennerl, K., et al. 2001, *A&A*, 365, L18
- Sturm, R. 2012, Dissertation, Technische Universität München, München
- Sunyaev, R. A. & Titarchuk, L. G. 1980, *A&A*, 86, 121
- Sunyaev, R. A. & Titarchuk, L. G. 1985, *A&A*, 143, 374
- Sutherland, W. & Saunders, W. 1992, *MNRAS*, 259, 413
- Tacconi, L. J., Neri, R., Genzel, R., et al. 2013, *ApJ*, 768, 74
- Thompson, A., Vaughan, D., et al. 2001, X-ray data booklet (Lawrence Berkeley National Laboratory, University of California Berkeley, CA)
- Toba, Y., Oyabu, S., Matsuhara, H., et al. 2014, *ApJ*, 788, 45
- Tozzi, P., Gilli, R., Mainieri, V., et al. 2006, *A&A*, 451, 457
- Treister, E. & Urry, C. M. 2005, *ApJ*, 630, 115
- Treister, E. & Urry, C. M. 2006, *ApJ*, 652, L79
- Treister, E., Urry, C. M., Chatzichristou, E., et al. 2004, *ApJ*, 616, 123
- Treister, E., Urry, C. M., Van Duyne, J., & GOODS AGN Team. 2005, *Mem. Soc. Astron. Italiana*, 76, 142
- Treister, E., Urry, C. M., & Virani, S. 2009, *ApJ*, 696, 110
- Tremaine, S., Gebhardt, K., Bender, R., et al. 2002, *ApJ*, 574, 740
- Trump, J. R., Impey, C. D., Elvis, M., et al. 2009, *ApJ*, 696, 1195
- Turner, M. J. L., Abbey, A., Arnaud, M., et al. 2001, *A&A*, 365, L27
- Turner, T. J., George, I. M., Nandra, K., & Mushotzky, R. F. 1997a, *ApJS*, 113, 23
- Turner, T. J., George, I. M., Nandra, K., & Mushotzky, R. F. 1997b, *ApJ*, 488, 164
- Turner, T. J. & Pounds, K. A. 1989, *MNRAS*, 240, 833
- Turnshek, D. A. 1988, in *QSO Absorption Lines: Probing the Universe*, ed. J. C. Blades, D. A. Turnshek, & C. A. Norman, 17
- Ueda, Y., Akiyama, M., Hasinger, G., Miyaji, T., & Watson, M. G. 2014, *ApJ*, 786, 104
- Ueda, Y., Akiyama, M., Ohta, K., & Miyaji, T. 2003, *ApJ*, 598, 886
- van der Marel, R. P. & van den Bosch, F. C. 1998, *AJ*, 116, 2220

- van Dyk, D. A., Connors, A., Kashyap, V. L., & Siemiginowska, A. 2001, *ApJ*, 548, 224
- Vasudevan, R. V., Brandt, W. N., Mushotzky, R. F., et al. 2013, *ApJ*, 763, 111
- Vasudevan, R. V. & Fabian, A. C. 2007, *MNRAS*, 381, 1235
- Vignali, C., Alexander, D. M., Gilli, R., & Pozzi, F. 2010, *MNRAS*, 404, 48
- Vignali, C., Mignoli, M., Gilli, R., et al. 2014, *A&A*, 571, A34
- Vito, F., Gilli, R., Vignali, C., et al. 2014, *MNRAS*, 445, 3557
- Wada, K. 2012, *ApJ*, 758, 66
- Wang, J.-M., Hu, C., Li, Y.-R., et al. 2009, *ApJ*, 697, L141
- Weisskopf, M. C., Tananbaum, H. D., Van Speybroeck, L. P., & O'Dell, S. L. 2000, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Vol. 4012, *X-Ray Optics, Instruments, and Missions III*, ed. J. E. Truemper & B. Aschenbach, 2–16
- Wilk, M. B. & Gnanadesikan, R. 1968, *Biometrika*, 55, pp. 1
- Woltjer, L. 1959, *ApJ*, 130, 38
- Worsley, M. A., Fabian, A. C., Bauer, F. E., et al. 2005, *MNRAS*, 357, 1281
- Xue, Y. Q., Luo, B., Brandt, W. N., et al. 2011, *ApJS*, 195, 10
- Yan, R. & Blanton, M. R. 2012, *ApJ*, 747, 61
- Yenko, B., Barger, A. J., Trouille, L., & Winter, L. M. 2009, *ApJ*, 698, 380
- York, D. G., Adelman, J., Anderson, Jr., J. E., et al. 2000, *AJ*, 120, 1579
- Zdziarski, A. A., Poutanen, J., & Johnson, W. N. 2000, *ApJ*, 542, 703
- Zel'dovich, Y. B. & Novikov, I. D. 1964, *Soviet Physics Doklady*, 9, 246